

“White Paper Report
from the
Biosignatures Standards of
Evidence Community Workshop”

Table of Contents

1. Are we alone?	3
1.1 Ongoing and future missions will enable the search for life beyond Earth	3
1.1.1 Missions on Solar System Terrestrial Worlds:	3
1.1.2 Missions on Solar System Ocean Worlds:	4
1.1.3 Exoplanetary Missions:	4
1.2 Past claims of life detection inform the processes required for biosignature assessment	4
1.3 The path to life detection	6
2. Towards A Generalized Framework for Biosignature Assessment	8
2.1 Overview of the generalized framework	8
2.2 General considerations when applying the framework	9
2.3 The Generalized Framework for Biosignature Assessment	10
Question 1: Have you detected an authentic signal?	11
Question 2: Have you adequately identified the signal?	12
Question 3: Are there abiotic sources for your detection?	13
Question 4: Is it likely that life would produce this expression in this environment?	15
Question 5: Are there independent lines of evidence to support a biological (or non-biological) explanation?	16
2.4 Future work needed to develop The Framework	17
3. Worked Examples of the Assessment Framework	18
3.1 Early Earth Worked Example: Detection and Characterization of Stromatolites ...	19
Question 1. Have you detected an authentic signal?	19
Question 2. Have you adequately identified the signal?	19
Question 3. Are there abiotic sources for your detection?	26
Question 4. Is it likely that life would produce this expression in this environment?	26
Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?	27
3.2 In situ Worked Example: Detection and Characterization of Organic Deposits on Early Earth and Mars	28
Question 1. Have you detected an authentic signal?	29
Question 2. Have you adequately identified the signal?	29
Question 3. Are there abiotic sources for your detection?	30
Question 4. Is it likely that life would produce this expression in this environment?	31

Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?	31
3.3 Remote Sensing Worked Example: Oxygen Detection on Exoplanets	32
Question 1. Have you detected an authentic signal?	33
Question 2. Have you adequately identified the signal?.....	33
Question 3. Are there abiotic sources for your detection?.....	33
Question 4. Is it likely that life would produce this expression in this environment?	35
Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?	35
3.4 An Agnostic Biosignature Worked Example: Distribution of Enantiomeric and Isotopic Composition of Chiral Molecules.....	36
Question 1. Have you detected an authentic signal?	38
Question 2. Have you adequately identified the signal?.....	39
Question 3. Are there abiotic sources for your detection?.....	39
Question 4. Is it likely that life would produce this expression in this environment?	41
Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?	42
4. Applications of the Framework to Missions.....	42
4.1 Utility of the Framework	42
4.2 Mission Sensitive Application of the Framework	42
4.3 Mission Life Cycle Considerations.....	43
4.4 Confidence and Statistical Inference	46
4.4.1 Biosignature Detection	46
4.4.2 Biosignature Assessment.....	46
5. Reporting Protocols for Life Detection Claims	48
5.1 Ideal goals for community biosignature verification and reporting.....	48
5.2 Current obstacles to ideal reporting protocol.....	50
5.3 Incentivizing the best possible verification and reporting process	52
5.4 How do we encourage and support collaborative verification?	53
5.5 Lessons learned.....	54
6: Findings and Future High Priority Research Avenues	56
6.1 Current understanding and findings of the workshop.....	56
6.2 Community Engagement Structures.....	60
References.....	64

1. Are we alone?

The detection of life beyond Earth is the principal goal of the field of astrobiology. However, the complexity of life itself, the scientific challenges of the required measurements, and the difficulty in conveying this complexity to others, means that it is now critical for the astrobiology community to develop standards of evidence and reporting for the detection of potential signs of life. A claim of life detection is a high-stakes scientific achievement that will garner enormous public interest, yet there is a high potential for any initial life detection claim to be ambiguous. Current and upcoming research efforts and missions aimed at detecting past and extant life will require a consensus framework to plan for, vet and discuss life detection claims (cf. Green et al., 2021) that avoids false positives and ensures clear communication with the scientific community and the public.

1.1 Ongoing and future missions will enable the search for life beyond Earth

In the past two decades, extraterrestrial life detection has become a more explicit and central goal for both planetary science and astronomy research objectives and missions. To have maximum applicability to current and future missions that search for life, any biosignature assessment framework needs to address discoveries from both missions that travel to the celestial bodies within our Solar System---that can collect orbital, fly-by or *in-situ* data, and missions that search for signatures of life on extrasolar planets via remote observation. This is not a trivial task, as both the signatures of life and the sources of false detections are strongly dependent on the source of the data itself, and the context of the environment in which it is obtained. In the coming decades, the developed framework will need to be applicable to searches for life on the following astrobiological targets, to which missions are either in progress or being planned.

1.1.1 Missions on Solar System Terrestrial Worlds:

Our terrestrial neighbors are of high astrobiological interest due to their relative accessibility, and the potential for comparison with Earth-based analog studies. The latter is key to how we learn to study life, and any framework for life detection must also apply to Earth-based studies, as well as extraterrestrial data. Mars has been a primary candidate for life beyond Earth, since the planet is likely to have possessed liquid water and broadly habitable surface conditions in its early history, and may even provide potential subsurface refugia for life today, including subsurface liquid water (e.g., Des Marais et al., 2008). The NASA Mars 2020 Perseverance rover is investigating evidence for past life in Mars's Jezero crater, and has recently collected samples for future transport back to Earth for detailed study including the search for biosignatures (Williford et al., 2018). The ExoMars program will also search for signs of past and present life on Mars, and includes a

currently operational orbiter investigating Martian atmospheric trace gases and their sources. This spacecraft will interface with the Rosalind Franklin rover (planned to launch in 2022) that will also search for water, variations in the geochemical environment, as well as organic molecules that could be indicative of past or present life on Mars. Venus is another terrestrial body of interest that may have been habitable in the past. Two NASA missions, VERITAS and DAVINCI, and ESA's EnVision mission are planned for launch in the next decade, and will provide new insight into the past and present environment of Venus, including the detailed composition of the current atmosphere and insight into potential false-positive biosignatures (e.g., Meadows, 2017).

1.1.2 Missions on Solar System Ocean Worlds:

The icy-moon ocean-worlds that orbit Jupiter and Saturn, such as Europa, Enceladus and Titan, have emerged as important targets in the search for life within the Solar System due to the availability of liquid water beneath their icy surface, and sources of energy (Hendrix et al., 2019). NASA's Cassini Saturn orbiter also detected the presence of organics on Enceladus and Titan (Postberg et al., 2018). A variety of missions have been selected and proposed to explicitly study the astrobiological potential of Europa (Europa Clipper: Pappalardo et al., 2017; Hand et al., 2017), Titan (Dragonfly: Turtle et al., 2019) and Enceladus (MacKenzie et al., 2021; Cable et al., 2016). Also, ESA has recently selected as a top priority the investigation of the habitability potential of the moons of the giant planets for future large-class missions in the timeframe 2035-2050.

1.1.3 Exoplanetary Missions:

In the coming decades, astrobiological studies outside our Solar System will gather data from remote observations using either ground-based telescopes and instruments or those in space. Characterizing the atmosphere and surface of potentially habitable planets and searching for signs of life via transmitted, reflected and emitted radiation is enabled in the next decade by current and near-term ground and space-based telescopes, including the James Webb Space Telescope (e.g., Krissansen-Totton et al., 2018), and the next generation of Extremely Large Telescopes (ELTs; e.g., Lopez-Morales et al., 2019). Future proposed missions that would enable more comprehensive searches for habitability and life on exoplanets include the NASA HabEx (Gaudi et al., 2020), LUVOR (The LUVOR Team, 2019) and Origins (Meixner et al., 2019) concepts, and ESA's LIFE mission concept (Quanz et al., 2021).

1.2 Past claims of life detection inform the processes required for biosignature assessment

While ongoing and future missions are poised to expand the search for life elsewhere, existing and past claims of possible biosignature detection within the Solar System have demonstrated the complex nature of the life detection process, and the community-wide efforts needed to assess these claims. The ensuing debates, within the scientific community and the public at large, also illustrate the necessity for an assessment framework to systematically increase the robustness

of the life detection claim being made, and a reporting protocol that could communicate the existing evidence and the assessment process more accurately, both within the field and to a wider audience.

One of the most famous claims for extant microbial life on Mars is from the Labeled Release (LR) experiments carried out by the NASA Viking Landers in the 1970s (Levin & Straat, 1976, 2016). The LR experiment tested for the presence of a metabolism in Mars regolith by injecting organics containing radioactive carbon-14 as a marker into a sample holder containing collected regolith and observing whether organics were consumed or transformed. While change was initially observed, a second gas injection produced no further labelled gas production and no organics were detected by Viking's Gas Chromatograph-Mass Spectrometer (GCMS) in the Martian soil or atmosphere. However, an abiotic origin for the LR result (Hecht et al., 2009; Glavin et al., 2013) is supported by results from both Viking's Gas Exchange (GeX) experiment and later discoveries by NASA's Phoenix and Mars Science Laboratory of the presence of oxychlorine compounds (CaO and Mg-perchlorates and chlorates) on the Martian surface. These compounds were found to react during the heating process of analysis, leading to oxidation and chlorination reactions that destroyed the organic molecules under investigation (Navarro-González et al., 2010). A biosignature detection framework applied to the LR result at the time of discovery could have more clearly indicated that while extant biology could be responsible for the gas evolution, an abiotic interpretation had not yet been ruled out, reducing confidence in the potential biosignature detected.

Claims for traces of past life in the martian meteorite ALH84001 is another example of a life detection claim that was later found to be more likely due to abiotic processes and characteristics. The McKay et al. (1996) team based their claims for the possible presence of life in the meteorite on multiple lines of evidence ranging from low temperature fluid infiltration of fractures in the cumulate volcanic rock, globular to chain-like physical structures in carbonate deposits identified as "nanobacteria", nanocrystals of magnetite identified as magnetosomes, and the presence of organic matter in the form of polycyclic aromatic hydrocarbons. However, at that time general understanding of the microbial world and how it is preserved in the rock record, as well as the potential for abiotic processes to create signatures that can imitate biogenic ones, was significantly limited. Today, this claim is discounted, although the very fact that we have a workshop on biosignatures hails back to this landmark paper.

More recently, claims of phosphine detection in the Venus atmosphere (Greaves et al., 2020), sparked a response that illustrated the multi-step process and necessity for community involvement in comprehensively assessing potential biosignatures. Although apparently detected in several observations with two telescopes, multiple sets of researchers reanalyzed the discovery data and questioned the significance of the detection (Snellen et al., 2020; Akins et al., 2021; Villaneueva et al., 2021; Thompson, 2021). Others searched for additional features

from phosphine at other wavelengths or with different instruments, in an attempt to corroborate the initial detection (Encrenaz et al., 2020; Trompet et al. 2020; Mogul et al., 2021). Modeling of the observed data suggested that the signal was more plausibly due to sulfur dioxide, a common gas in Venus' atmosphere which also absorbs near the same frequency position, rather than PH_3 (Lincowski et al., 2021; Akins et al., 2021; Villaneueva et al., 2021), although this too is now contested (Greaves et al., 2021). While the discoverers performed an exhaustive survey of potential abiotic production mechanisms but could not identify a likely source (Greaves et al., 2020; Bains et al., 2021), others worked to find plausible abiotic formation mechanisms for PH_3 in the Venus atmosphere (Truong & Lunine 2021 Omran et al., 2021). This rapidly evolving scientific story illustrates the progressive way in which science combines a series of results to more accurately assess the detection of possible biosignatures.

Claims of signs of life in the early rock record similarly illustrate the need for consistent criteria in defining biological origin, as well as the importance of contamination evaluation and the enhanced rigor enabled when multiple teams engage in assessment of claims of biosignature detection. In one example, Schopf (1993) described "microfossil-like objects", which he interpreted as eleven different microbial taxa in a ~3.5 Ga sedimentary chert rock. These structures were reinterpreted by Brasier et al. (2002) as abiotic organic matter in a hydrothermal setting. The debate has continued for over 30 years, with Rouillard et al. (2021) noting that the criteria used for biological origin was inconsistent within the community, causing the fossil-like structures to sometimes be presented as evidence for biological activity, and other times as abiotic. This example underscores the need for community-wide adoption of a standard set of criteria for assessing potential preserved biosignatures. In another example, Brocks et al. (1999) identified hopane and sterane organic compounds in Archean sedimentary rocks, which can be indicative of bacterial and eukaryotic cells, respectively, having used the best protocols at the time to rule out contamination. This research led to the interpretation that eukaryotes flourished 300 million years prior to the oxidation of the atmosphere, a view that was supported for almost a decade until Rasmussen et al. (2008) showed that while organic matter was indigenous to the rock, the hopane and sterane biosignatures were likely, in fact, the result of contamination. This was confirmed by French et al. (2015) who conducted a drilling campaign that was specifically designed to further minimize contamination, followed by analysis in multiple laboratories utilizing clean procedures. This example also demonstrates the importance of a rigorous and robust multi-laboratory approach in increasing the confidence in biosignature detection.

1.3 The path to life detection

Given the examples above, it is clear that the detection of extraterrestrial life is unlikely to be instantaneous or unambiguous, and yet it is a high stakes scientific achievement that will garner enormous public interest. Even when examining evidence of early life on Earth under the most ideal analytical conditions, detecting and interpreting biosignatures is extremely challenging, requiring multiple

measurement techniques and contextual information, sometimes developed over decades. For the arguably more ambitious goal of detecting extraterrestrial life, we will likely embark on even longer journeys involving a series of measurements and missions, each progressively increasing our confidence in the interpretation of our findings, and helping to provide key aspects of biosignature detection, confirmation and assessment. It is crucial for our field to clearly convey the complex, multi-measurement assessment strategy necessary for biosignature detection. Such a framework will allow us to articulate the inherent value of a measurement or mission in increasing our confidence in life detection, and how it advances us on the path to life detection, even when that mission is not returning definitive evidence of life itself.

With claims of possible life detection becoming more frequent, it is critical to discuss as a community how to agree upon a universal scientific framework to best vet and support discoveries, and to report and convey information on a topic that is inherently complex and interdisciplinary. Green et al. (2021) articulated one such example framework for biosignature assessment and encouraged the astrobiology community to refine and develop their own. The Biosignature Standards of Evidence (SoE) Workshop took up that charge, and additionally discussed how to best communicate claims of life detection to our scientific colleagues and to the public.

In the remainder of this document, we present a community-developed, universal framework that can be used to improve assessment protocols for biosignature detection claims, contextualize a measurement or mission result, and more accurately communicate confidence in results related to life detection. Construction of such a framework is not trivial, as different fields will have inherently different ways to assess their confidence in biosignature detection and assessment, and there is an inherent tension between a rubric that is simple to communicate and use across a wide range of life detection scenarios, but nuanced enough to handle the complex physical and chemical phenomena measured using a wide range of techniques. In particular, investigations in our own Solar System generally allow for the collection of samples or analysis *in situ*, yet the search for life on exoplanets depends on telescopic measurements from great distances. Reconciling the standards of life detection between these two communities is complex, but crucial. With this in mind, the framework described through this paper was developed via a massive interdisciplinary collaboration from the different fields within the astrobiology community, and the hope is this will be continuously improved upon and developed over time as we progress in our search for life.

In this white paper, Section 2 outlines the structure of the framework itself. Section 3 provides historical and hypothetical examples of claims of possible biological activity to illustrate how the framework might work in practice for a range of targets and measurement techniques. Section 4 explores deeper applications of the framework to multiple phases of mission development and lifetime. Section 5 describes a protocol for reporting biosignature assessment results. Section 6

outlines possible pathways of research that would help to fortify the aims of this framework.

2. Towards A Generalized Framework for Biosignature Assessment

In this section we describe the first steps in developing a generalized framework for biosignature assessment. As motivated by Section 1, to scientifically assess and accurately communicate likelihood in claims of life detection, there is a need for a community framework that identifies key questions and provides a generalized description of the scientific processes and evidence needed for this assessment. We use the term '*generalized*' here to denote that the ideal framework is sufficiently broad and process-based that it can be useful when applied to life detection claims made using a variety of measurement techniques, including both *in situ* and remote-sensing, for a variety of targets within and beyond our Solar System, and for extinct and extant life. This framework would be applied either once a claim of biosignature detection is made, or proactively, to justify the efficacy of measurements and theoretical work in enhancing future life detection efforts. In this section, we outline the assessment framework structure that was developed and refined at the Biosignature Standards of Evidence workshop, and describe in more detail the generalized scientific questions to be answered in each step.

2.1 Overview of the generalized framework

The assessment framework comprises five Questions clustered into two broad components (Figure 2.1): Level 1, which assesses our confidence in the detection and identification of a potential biosignature, and Level 2, which assesses whether or not the potential biosignature is more or less likely to be due to life, especially in the context of the environment in which it is found. More specifically, Level 1 of the framework encourages those who discover a potential biosignature to address Questions 1 and 2, which are: “**Have you detected an authentic signal?**”, and “**Have you adequately identified the signal?**”. Level 2 moves from detection and identification, to assessment of the likelihood that the signal is due to life and includes Questions 3, 4, and 5: “**Are there abiotic sources for your detection?**”; “**Is it likely that life would produce this expression in this environment?**”; and “**Are there independent lines of evidence to support a biological (or non-biological) explanation?**”.

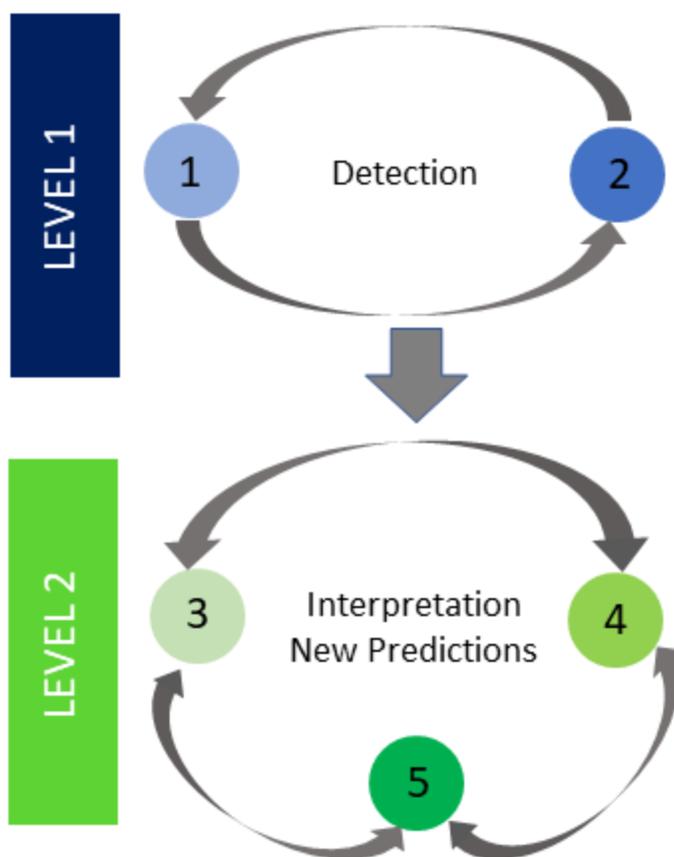


Figure 2.1: Schematic of the Community Biosignature Assessment Framework. Level 1 encompasses Questions 1 and 2 to be addressed to increase certainty in potential biosignature detection and identification, and Level 2 includes Questions 3-5 centered on assessment and discrimination between abiotic and biological processes as the source of the potential biosignature, and may generate new predictions needed to enhance confidence in life detection.

2.2 General considerations when applying the framework

While the framework questions are presented in order, there was a strong sense at the workshop that the application of the steps do not need to be in a particular order, and in some cases may be difficult to implement linearly. Within the levels, measurements and processes may be cyclical or iterative, as a given measurement or modeling result might simultaneously address several questions, and further measurements or other research to address subsequent questions may in fact reduce, rather than enhance, the degree of certainty in a given interpretation. However, to minimize impact on the broader community from a biosignature detection claim, it is considered desirable to have high confidence in Level 1 (the biosignature candidate detection and identification) prior to moving on to Level 2, where significant community efforts will be needed to assess whether a biological source is a likely explanation for the potential biosignature.

The assessment framework is also designed to be generalized, and therefore applicable to a broad range of biosignature types and measurement techniques. While the specific application of the framework and the assessment of certainty in a given measurement or interpretation will vary as a function of measurement technique, target and biosignature sought (see Section 3 for more detailed field-specific examples), the questions that we ask ourselves as astrobiologists---and the processes we must perform to enhance scientific robustness---are more universal. For example, confidence in a biosignature detection is deeply dependent on an understanding of instrumental systematics and biases, a process common to all, but for which the specifics vary depending on instrument and field. Similarly, confirmation via multiple measurements or additional instruments was considered more robust than a single measurement alone, but which instruments could be brought to bear was very field and biosignature specific.

Moreover, in grasping and applying this framework, challenges can arise in the different terminology used by different fields for similar concepts, or for different processes that have similar impacts on our ability to detect or assess biosignatures. For example, both coelution of compounds in a chromatographic sample, and blending of lines or bands from two different molecular absorbers in the spectrum of a planet, make the unique identification of a potential biosignature more difficult. Even though the techniques and targeted biosignatures are extremely different, the process of disambiguation of these measurements is a universal step that would be required to increase the robustness of biosignature detection and assessment. Consequently, it is expected that although the biosignature assessment framework contains generalized questions and processes, its application will be different and adaptable to the needs of a given technique and its targeted biosignature.

Other discussions around applying a more generalized framework also centered on the recognition that addressing certain questions may be more challenging, and require a different threshold for confidence depending on the investigation type, whether that be *in situ* or remote, or ancient vs. extant life. To help maintain a generalized framework, it was considered valuable that these types of measurements should be assessed on the same confidence scale, even if very different certainties were considered achievable due to inherent limitations in the measurement technique or data sought. Discussions were initiated on the difficulty of assigning numerical confidence to a given question, especially across fields, and whether a more simplistic three-level scale might be more appropriate. Combining different confidence levels for different questions into a single confidence value was also discussed as a possibility, as was at least marking when a given study had cleared Level 1 to the community's satisfaction.

2.3 The Generalized Framework for Biosignature Assessment

Below, we describe the framework questions in detail, providing a plain language summary of the intent of the question, the motivation to address the question, and

known techniques and scientific strategies for how we might best address the question.

Question 1: Have you detected an authentic signal?

Do you have a real, statistically significant signal? Have you ruled out artifacts from the measurement, pre-processing and/or analysis process that might mimic a real signal?

A claim of biosignature detection begins with an observation or a measurement of a particular phenomenon, from structures observed in *in situ* collection of sample data, to a remotely sensed molecular absorption. The first assessment question is then “is this a real signal?”. Answering this question first is critical, because if there is a low probability that the signal is real, then biosignature assessment does not need to proceed to subsequent steps. Reasons for which the signal may not be real include the injection and/or misinterpretation of noise or contamination due to an instrumental, measurement technique or processing artifact. For example, in planetary spectroscopy, reflections in the instrument or other noise sources may produce an apparent signal, where in reality none exists. In the laboratory, extracting organics with hot liquid solvents may generate molecules that did not exist in the original sample. The key guiding question then becomes, what unknowns, noise, measurement technique, instrumental or analysis limitations must be addressed to consider the detection legitimate?

There are several procedures and techniques that can be used to increase confidence that a given biosignature detection represents a real signal. The signal characteristics should be of adequate quality, including signal-to-noise ratio (SNR) and measurement frequency. *A priori* knowledge of instrument performance and limitations through validation is an important factor, as it helps to identify previously known spurious signals from the instrument, and also helps to determine if the reported signal is within the instrument’s expected sensitivity and performance. Calibration blanks or other reference measurements used at the time of, or in conjunction with the measurement, can be used to rule out spurious instrumental or measurement system signals. These validations and calibrations, as well as other key aspects of the way samples are taken, or the way the phenomenon is measured, should be used to estimate the uncertainties and the corresponding detection significance using standard field-specific techniques. Confidence in the signal can also be increased via multiple detections on the same sample, potentially with multiple instruments that provide convergent data, and if possible on different samples. Similarly, reanalysis of the sample or detection data by multiple independent teams can be used to strengthen confidence in the detection. Data processing and statistical analysis techniques, such as resampling via bootstrapping or using other perturbations can also be brought to bear to test the stability of the signal, and estimate true confidence intervals for the likely range of values associated with the detection. The repetition of the detection with temporally-spaced samples may also increase confidence that it was not due to

transient noise, as long as the phenomenon is expected to persist over time, and that any spurious instrumental signals or contamination is not.

Question 2: Have you adequately identified the signal?

Have you adequately ruled out other potential sources for this signal? For example, have you ruled out contamination in the environment, or other real phenomena that could produce a similar signal?

Having established a high degree of confidence that the potential biosignature signal is an observed phenomenon in the sample or planetary target, and not due to instrumental, measurement technique or processing artifacts, Question 2 then addresses whether the observed signal has been adequately identified, and attributed to the correct phenomenon. Question 2 therefore focuses more on issues of misidentification of the biosignature due to contamination via phenomena that do exist elsewhere in the environment or the sample, but are not the target biosignature sought. Like Question 1, addressing this Question increases the likelihood that a potential biosignature has indeed been detected, prior to the more comprehensive and interdisciplinary effort outlined in Questions 3-5 that are needed to interpret that biosignature detection. In some fields, addressing Questions 1 and 2 may need to be done, or naturally occur, simultaneously, rather than as separate steps. Specific examples of possible phenomena that may confuse biosignature identification include misidentification or coelution of spectral or chromatograph lines from multiple molecules absorbing at the same position, or contamination of rock samples with non-indigenous organics. The key guiding question becomes: can you be reasonably sure that the observable signal is representative of the molecule, process, or other line of evidence to which it is attributed, and for the sample or target under consideration, and is not due to another phenomenon?

There are multiple techniques and procedures that could be used to enhance confidence in the correct identification of the signal. Understanding the limits of detection or resolution for the observed signal can be a first step toward determining whether a misidentification is likely. Other steps to confirm identification can include comparing the measurement with known measurement-specific standards or libraries, which can also be used to reveal potential confusion with other phenomena. Steps can then be taken to account for contamination of the sample or other interactions with the environment. Multiple samples can be taken and compared to look for consistency. Similarly, comparisons with procedural or instrument “blanks” can reveal contamination. Sample history, including determination of syngeneity (that the age of the rock and any putative biological material within is equivalent; Oehler & Cady, 2014) and spatial relationships to structures in the sample, can help discriminate contamination by more recent biological material. Careful site selection for sampling, and recording of sampling context can also help with contamination control and identification. Obtaining additional, complementary measurements and contextual clues to break any potential degeneracies can include confirming the identification with multiple

instruments and/or features. For example, searching for additional absorption features from a molecule in spectra taken with different instruments at different wavelengths, or using multiple techniques on a single rock sample to provide corroborating isotopic, compositional and spatial information to identify embedded organics. Finally, contextual clues from the environment may provide further support for the identification by showing that the detection is consistent with other physical, chemical, and/or geological parameters detected or characterized at the site.

Question 3: Are there abiotic sources for your detection?

Is it likely that there is a current or past environmental process, other than life, that could be producing this signal? Have you ruled out potential false positives for the biosignature?

Once the detection of a signal has been confirmed and possible contamination has been ruled out (Questions 1 and 2), the assessment framework moves towards the steps needed to determine whether or not the signal is due to a biological process (Questions 3-5). The first of these steps (Question 3) is to determine if there are any plausible abiotic sources of the signal, and if so, to attempt to rule them out. This is a critical first step in assessing the biogenicity of a signal as it conservatively assumes that the signal is not due to life. If a convincing abiotic explanation is found, it can reduce the probability that a true biosignature has been observed. However, as each alternative abiotic explanation is proposed and then potentially ruled out, it can also strengthen the case---although not prove---that the biosignature may be due to life. While a biosignature by definition reveals life's impact on its environment, our interpretation of the biogenicity (or abiogenicity) of a detection will in turn strongly depend on the (paleo)environmental context. A recognized biosignature in one environment may not directly translate to another, and abiotic planetary processes in a different environment may produce "false positives" for the biosignature. Examples of proposed false positives for putative biosignatures include: modeling of the production of atmospheric O₂ via ocean loss and photochemistry, rather than photosynthesis or other biological processes (cf. Meadows, 2018, and references therein); laboratory generation of prebiotic compounds and synthetic organic chemistry (Barge et al., 2021); observations of abiotically sourced molecules detected at hydrothermal vents (Ménez et al., 2018); measurements of amino acids and other structurally intricate organic compounds in meteorites (Burton et al., 2012; Glavin et al., 2020); and measurements of aromatic organics generated during testing of Martian samples by the Mars Science Laboratory (Curiosity Rover) (Szopa et al., 2020).

Key challenges in identifying abiotic processes for *in situ* and early Earth studies include: limitations to our understanding of how geochemical systems evolve over time; the preservation potential of abiotic signatures; the historical focus on simulating prebiotic chemistry in early Earth environments rather than conditions relevant to other worlds; technological constraints in characterizing organic content on other planetary bodies; and, of course, the potential sheer enormity of

environments and processes that may produce abiotic signals. For exoplanet studies, the diversity of potential planetary environments and processes is likely even larger, and must be explored solely via modeling until observational data on exoplanet environments becomes available. However, our understanding of terrestrial exoplanet environments can be informed by geological and atmospheric planetary processes from planets in our Solar System (Kane et al., 2021), as well as stellar/planetary/planetary system interactions.

However, despite these challenges, there are some potential best practices. Published methods describing abiotic sources for the potential biosignature, or chemically similar compounds, should be sought. This includes whether or not a molecule is detected to high confidence in astromaterials or the interstellar medium, or has been generated in laboratory analog or prebiotic chemistry studies. If an abiotic synthesis mechanism is known, then other molecules or signals that might be part of that mechanism should be identified and sought to help confirm or exclude abiotic generation. The presence of mineral or metal catalysts should be noted for their potential to drive additional chemistry. The stability of the molecule under the conditions in which it is found should be considered, including whether its abundance changes over time. An active flux could be an abiotic release mechanism, possibly even sourced from a preserved residual component of extinct life. Alternatively, when looking for signs of life in remote-sensing data, an active, potentially changing flux of a molecule (e.g., seasonally) may be biological and so requires further investigation. The stability of the molecular isomers, enantiomers, and homologs should also be considered and can aid in the evaluation of the detection being sourced via abiotic mechanisms. In particular, whether abiotic sources consistently lead to the production of homologs and isomers of the proposed biosignature should be assessed. If so, their absence (if the technique was able to detect them) would build a stronger case against an abiotic origin; if homologs or isomers could not be detected or resolved with the given detection method, this should be noted. In addition, the stability of the putative biosignature under the conditions it was found should be considered. Long term resilience or “preservation potential” of a biosignature is a key parameter (along with origination/plausibility and concentration) for a reliable biosignature interpretation (cf. Westall et al., 2018).

Because abiotic signatures may have been generated earlier in the environmental lifetime, and been preserved, mechanisms that may have operated under conditions different to the present day environment must also be considered. For morphological signals, considerations of whether or not the signal is widespread, and the degree of degradation, will be important considerations for assessing whether an abiotic source for the structure is more likely.

Finally, for consideration of abiotic processes in assessing the biosignature, we must evaluate the state of community knowledge pertaining to those processes or classes of processes. If the only proposed abiotic model or mechanisms are poorly characterized, or if the scope of possible abiotic explanations is known to be poorly

explored, it suggests we cannot adequately reject abiotic mechanisms. Any detection claim should include the identification of areas of further study that would be required to identify or preclude possible abiotic sources. This information should be considered in combination with the analysis in Question 4, on the likelihood of the life hypothesis.

Question 4: Is it likely that life would produce this expression in this environment?

Given what we know about the likely environment that an organism is operating in, or would have operated in, does it make physical and chemical sense that life would produce this potential biosignature?

Critically, life detection claims should not only rest on ruling out abiotic sources, but also on the viability of biological explanations. A claim of extraterrestrial life detection is extraordinary, and this hypothesis should be rigorously tested in its own right, and not simply considered the most likely explanation once all known potential abiotic mechanisms have been sought and excluded. For a signal to be a biosignature, it must be consistent with the products or activity of life whose nature can be supported by the (paleo)environment in question. Addressing Question 4---and making the case for the likelihood that life could have produced the signal in the environment under consideration---therefore complements efforts in addressing potential abiotic sources in Question 3, and both activities work towards attempting to support or rule out a biological explanation for a biosignature. Key considerations relevant to answering Question 4 include determining whether or not the environment in which the candidate biosignature has been found is currently habitable, and whether it contains the resources and physicochemical conditions required to support life, or whether it could have potentially supported life in the past. A justification also needs to be developed as to why life would make this signal in the observed environment. This justification may involve comparison with the known behavior of Earth organisms in similar environments, or may require the development of a model of life that is consistent with what is known of the non-Earth-like environment. Important considerations in developing such a model include whether or not the production of the biosignature is physically or chemically plausible in the environment. For a biosignature that may have been generated by ancient life, the likelihood of signal preservation also needs to be considered.

A habitability assessment is a key component of determining whether it is likely that life can produce (or could have produced) a biosignature in the environment under study. In many cases, an initial habitability assessment will already have been performed for site or target selection before the biosignature measurement was made. These assessments are necessarily bounded by what we know about the limits to life on Earth. Where these limits have been exceeded in a specific environmental parameter, the biosignature could, depending on context, be rejected. Alternatively, further analysis could be undertaken to discover if that particular parameter measurement is reliable, or could plausibly be accommodated by biochemistry. The detection of a potential biosignature may

spur further environmental characterization on multiple spatial and temporal scales, in order to improve our ability to identify false positive mechanisms (see Question 3) and to assess the likelihood that life could exist in that environment.

Attempts to justify whether life might produce the detected signal in the measured environment can also consider if an assumed model of life's impact on its environment is physically and chemically consistent with the phenomena observed, and consistent with what is known of the environment. Current bold goals include progress towards models that can determine if it makes environmentally contextual sense that life produced the observed phenomenon can be made without a definitive definition of life. If life on Earth is known to produce similar signals in analog environments, then a direct comparison can be made. However, if the extraterrestrial environment has no Earth analog, a model of how life might function in that non-Earth environment, and plausibly produce the detected potential biosignature, may need to be proposed. For example, if the potential biosignature is proposed to be a metabolite, then assessment of whether life has a thermodynamic impetus or benefit in creating such a product in that environment would need to be developed.

There are a number of ways that the plausibility of life producing an environmental phenomenon could be assessed. For example, plausibility arguments can potentially be made by looking at the abundance or inferred flux of the potential biosignature and checking whether it is consistent with estimates of the free energy or nutrient availability in that environment, and whether it is out of equilibrium with its environment. For biosignatures that are characteristic physical patterns, molecular distributions, or complex signals or molecules, probabilistic arguments can be made for biological processes being more likely than abiotic ones to produce these features. For biosignatures of fossilized or remnant life, a further plausibility consideration is whether it is likely for the observed signal to have been preserved over geologic time in this environment, or be more likely due to forward contamination. Beyond determining if the potential biosignature "makes sense" given the environment in which it is observed, broader considerations of the plausibility that life would produce this signature include considerations of the probability that life emerged on the body in question. However, this requires the incredibly challenging acquisition of constraints on the environment in which life likely evolved, which may be very different to the environment under study. This ancient context is tied to the planet's evolutionary history, which may be somewhat knowable (e.g., for Mars) or extremely ill-defined (e.g., for exoplanets).

Question 5: Are there independent lines of evidence to support a biological (or non-biological) explanation?

Are there other measurements that provide additional evidence, or allow you to predict and execute follow-on experiments, that will help discriminate between the life or non-life hypotheses?

Although both the hypotheses of non-life and life will have already been tested in Questions 3 and 4, it is likely that the question of the presence or absence of life will still be unresolved and will require ongoing iteration. Addressing Question 5 is intended to provide additional rigor to the investigation by encouraging the further acquisition of multiple independent lines of evidence, and by the prediction and testing of hypotheses in support of either the biological or non-biological explanation for the observed potential biosignature. Developing theories for how life may be generating the signal, including characterizing the environment in which it is found, requires constantly improving currently incomplete models of geophysics, geochemistry, molecular biology, taphonomy, and many other disciplines, and so single lines of evidence are anticipated to be inadequate for assessing purported biosignature detection reports. If multiple lines of evidence have not been obtained, then follow-up experiments are likely required. A gold standard of any life detection study is to make predictions about what other features or phenomena should be observed in order to be consistent with the conceptual and theoretical framework that is used to assess Questions 3 and 4 (see Green et al., 2021).

Further assessment of biosignature interpretation could include historical research in the literature, reanalysis of existing data by multiple teams, and the development of new follow-on experiments. A thorough check should be made to determine if previously published research supports or contradicts the hypothesis to be tested, and multiple existing studies could be analyzed to better understand reproducibility (Botvinnik-Nezer et al., 2020).

Authors should make their methods and data readily available so that other groups can replicate results. When possible, independent labs and research teams should confirm potential biosignature observations. It is not necessary that results be replicated by multiple teams before publication of the discovery, but the strength of conclusions should be moderated, if not replicated prior to publication. It is important to note that multiple lines of evidence from different labs and/or research groups do not *guarantee* that the observation is a true, valid biosignature. However, given the importance of life detection claims, independent replication of observations and data analysis is crucial, since higher credibility will be afforded to multiple observations of the same biosignature, and/or a collection of potential biosignatures.

2.4 Future work needed to develop The Framework

The Biosignatures Standards of Evidence Workshop allowed us to make significant inroads in developing the above community consensus framework. However, further work, including potentially another workshop, will be needed to determine additional and field-specific details on how we can best assign confidence levels based on the data available, and whether a numerical, color, or other scale is the most useful in communicating degrees of certainty about the veracity of a given biosignature claim. Another potential avenue for further community discussion is the possible usefulness of combining responses to the

questions into a single numerical or color ranking that conveys certainty in the current assessment of the biosignature claim. Such a task may be quite difficult to implement, however, and may need to depend on an external body for biosignature assessment.

Future work may also include a discussion on how best to use the framework in conjunction with the NASA Life Detection Knowledge Base (LDKB), which provides a community-driven assessment of potential biosignatures to be sought, and the potential for false positives and false negatives associated with specific measurements. The LDKB can serve as a central repository for community knowledge on how well studied a particular biosignature is, including its false positives and negatives, to aid in the selection of target biosignatures. The LDKB could also be used in conjunction with the biosignature Assessment Framework to inform the utility of specific life detection measurements in confirming the presence of life or revealing potential abiotic sources and potentially help with mission concept development (cf. Section 4).

Finally, the Framework can be used to inspire the laboratory, field and modeling communities to work towards comprehensively exploring plausible abiotic sources of a particular signal, determining the relevant follow-up signals necessary to distinguish abiotic cases, and ensuring that missions are designed with the capability of measuring these additional signals to confidently rule out all known false positive scenarios. This also allows appropriate resources to be allocated to any particular biosignature within a mission (e.g., observing time to achieve desired signal-to-noise in a given spectral band, or measurement of more isotopic ratios to increase precision), and identifies which mission aspects might benefit from additional resource allocation.

3. Worked Examples of the Assessment Framework

The five steps of the Assessment Framework were discussed at length over two days during the Standards of Evidence workshop, both within groups that focused on particular biosignature classes, disciplinary specialties, analytical targets, or instrumentation, and within groups that were purposefully randomized. All of these groups were asked to address two particular questions: Can the five steps of the Assessment Framework be mapped to the needs of your specialty? and Are there key missing aspects that would affect applicability? To summarize these activities, we have prepared four worked examples to demonstrate how the Assessment Framework can be used. We include an example from paleobiology, a discipline that has long grappled with distinguishing signs of life in the geologic record, an example from the analysis of organic matter on Mars to demonstrate an in situ application, a spectroscopic example that demonstrates utility for remote sensing applications, and an example that uses an “agnostic” lens on a possible biosignature without relying on known life on Earth for comparison.

3.1 Early Earth Worked Example: Detection and Characterization of Stromatolites

Stromatolites are laminated sedimentary structures that form at the sediment water interface in freshwater, marine, and evaporitic environments under the influence of microbial growth or metabolic activity. This definition aligns with Grey and Awramik (2020), who consider stromatolites to be a subset of microbialite (for definitions of other types of microbialite see also: Riding, 2011; Bosak et al., 2013). Stromatolites occur widely in the Precambrian rock record and, because they indicate a history of microorganisms modifying their immediate environment, are accepted as evidence of some of the earliest life on Earth (e.g., Allwood et al., 2006; Schopf et al., 2007). There are many stromatolite and other microbialite discoveries from the early Archean (e.g., Lowe, 1980; Walter et al., 1980; Byerly et al., 1986; Buick et al., 1981, 1995; Westall et al., 2006; Allwood et al., 2009, 2010; Wacey, 2010; Nutman et al., 2016; Djokic et al., 2017; Hickman-Lewis et al., 2018; Homann et al., 2018), a number of which have subsequently been questioned (e.g., Lowe, 1994; Schopf et al., 2006; Bosak et al., 2013; Allwood et al., 2018; Javaux 2019; Lepot, 2020, and references therein). The Assessment Framework could provide a more universal basis on which the validity of these discoveries can be examined. It can be difficult to assess the biogenicity of any one purported stromatolite, especially if it has been altered through diagenesis and/or metamorphism (e.g., Nutman et al., 2016; Allwood et al., 2018). It is also recognized that abiotic processes can mimic the broad morphology of stromatolites (e.g., Grotzinger & Rothman, 1996; Chafetz & Guidry, 1999; McLoughlin et al., 2008), which can further complicate identification. Though the biological information preserved from early Earth is often patchy and/or overprinted, the biogenicity of a purported stromatolite can usually be distinguished using several lines of evidence across multiple scales (*sensu* Question 5; Lepot, 2020, and references therein) and by comparison to modern examples and growth models and experiments (e.g., Batchelor et al., 2003; Petroff et al., 2010; Hickman-Lewis et al., 2019).

Question 1. Have you detected an authentic signal?

Laminated sedimentary structures of possible biogenic origin are initially identified in the field from visual observations, based on stratigraphic occurrence, bulk mineralogical fluctuations and the macroscopic morphology of the structure and its relationship with surrounding layers (Grey & Awramik, 2020).

Question 2. Have you adequately identified the signal?

While in the field, assess whether the laminated structures are primary and indigenous to the rock by observing them from a hierarchy of levels (spanning mega-, macro-, meso-, and microscopic scales), focusing on their various components (e.g., layers, textures, grains, etc.; Grey & Awramik, 2020; Fig. 1). Investigate whether they are only present on the surface of the rock (which may be indicative of a mineral deposit or weathering expression), or are embedded in

(i.e., are a part of) the rock. Also consider how much the structures may have been impacted or even created by secondary geological processes (e.g., overprinting, veining, mineral replacement, fluid flow). This may require sampling and further analysis in the laboratory (e.g., thin section petrography, electron microscopy, 3D imaging of morphology, geochemical analysis). Replicate field observations and/or analyses may also be required.

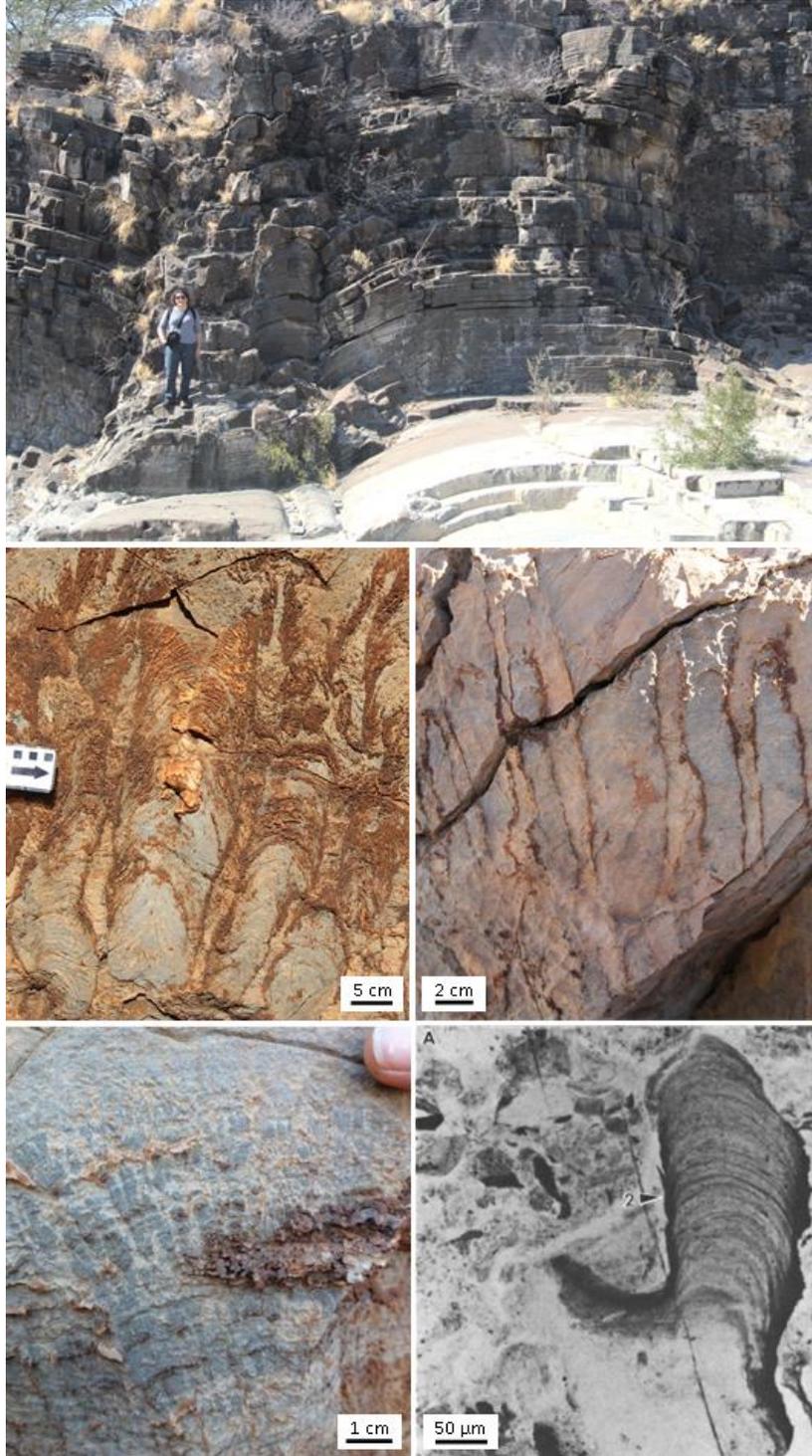


Figure 1: *Stromatolites can vary dramatically in size, from metres (A), to centimetres (B, C), to millimetres (D), to micrometres (E) in width. This makes it important to assess any laminated structure (i.e., purported stromatolite) at a range of scales. A) Broad, domal stromatolites from c. 2.6 Ga Reivilo Formation, Boetsap, South Africa. B), C) Columnar stromatolites from c. 2.4 Ga Turee Creek Group, Western Australia. D) Microdigitate stromatolites from c. 2.4 Ga Turee Creek Group, Western Australia. E) Micron-sized stromatolite from c. 2.5 Ga Monte Christo Formation,*

Chuniespoort Group, South Africa. Photo credit: A, AD Czaja; B and D, Barlow et al. (2016); C, EV Barlow; E, Lanier et al. (1986).

Additionally, one can investigate the interplay between the laminated structure and the surrounding sediments; can you distinguish between the two? In cross section, many stromatolites exhibit vertical relief above the corresponding layer of sediment, reflecting the interaction between processes that are biologically influenced and those that are purely sedimentological/physical (e.g., Van Kranendonk, 2011; Fig. 1).

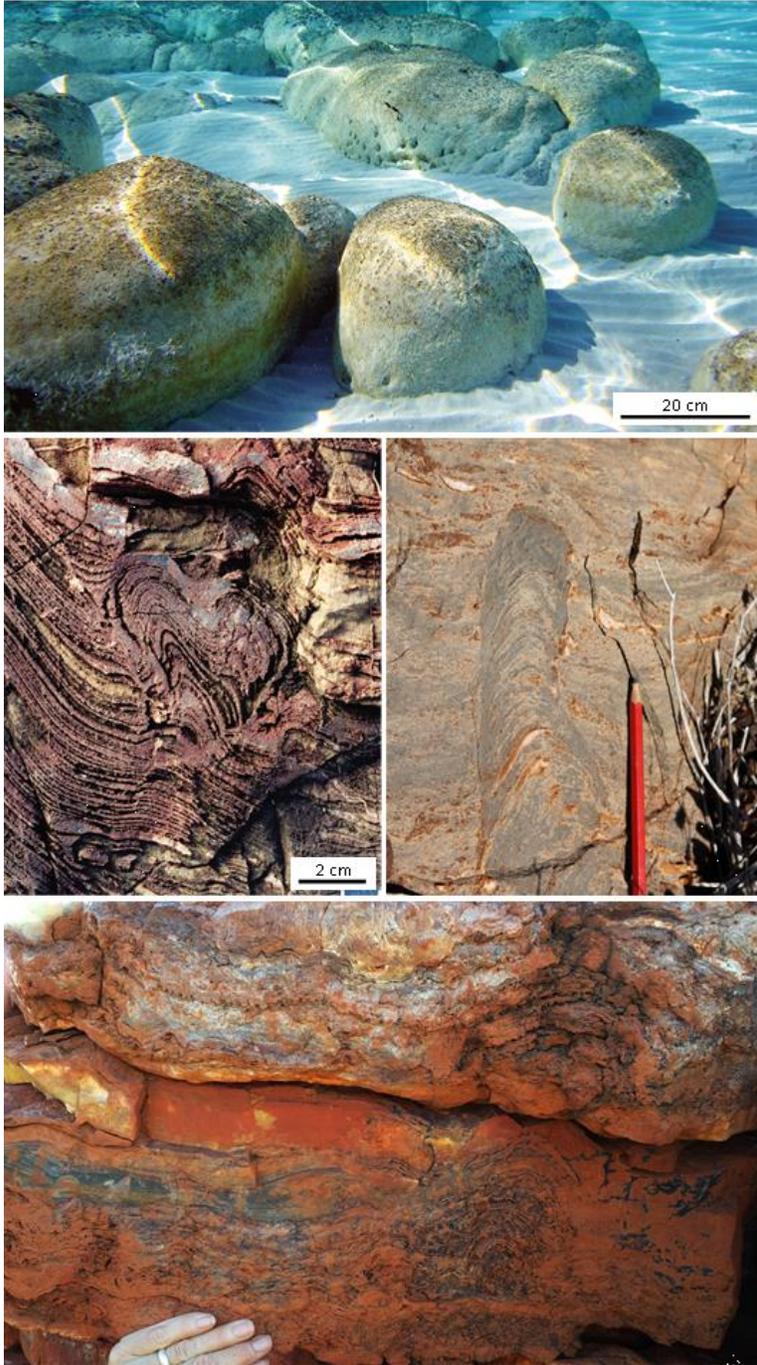


Figure 2: While forming, stromatolites generally stick out above the sediment on the seafloor (A). This is then translated in the rock record as a perpendicular structure that is distinct from the adjacent layers (B-D). A) Hemispherical stromatolites in Hamelin Pool, Shark Bay, Western Australia. B) Branching columnar stromatolite (“Mickey Mouse ears”) from c. 3.43 Ga Trendall Locality, Strelley Pool Formation, Western Australia. C) Columnar stromatolite from c. 2.4 Ga Turee Creek Group, Western Australia. D) Columnar stromatolite from c. 3.48 Ga Buick Locality, Dresser Formation, Western Australia. Photo credit: A, SM Awramik (in Grey & Awramik, 2020); B, K Grey (in Grey & Awramik, 2020); C, EV Barlow; D, MJ Van Kranendonk.

Next, determine whether the laminated structures are both vertically and horizontally extensive throughout the stratigraphy. Stromatolites commonly grow in clusters, or bioherms, with many individual specimens preserved in close proximity to one another (Fig. 3; Allwood et al., 2006). If stromatolites are present, then other types of microbialites (e.g., thrombolites, dendrolites, leiolites, or microbially-induced sedimentary structures (MISS)) may occur in close proximity, and there may even be a transition from one type of microbialite to another within a single structure (Grey & Awramik, 2020). As such, multiple observations of the same, or similar, type of laminated structure within an outcrop can increase detection confidence. Biogenic stromatolites can also feature micromorphologies incompatible with abiotic morphogenesis. For example: predictable non-isopachous laminations that thicken over the crests of domes due to phototrophic behaviour (Pope et al., 2000); overhangs that maintain shape by virtue of the cohesive properties of microbial sediment (Cuerno et al., 2012); and supra-lamina complexities that develop in response to environmental stressors (Hickman-Lewis et al., 2019).

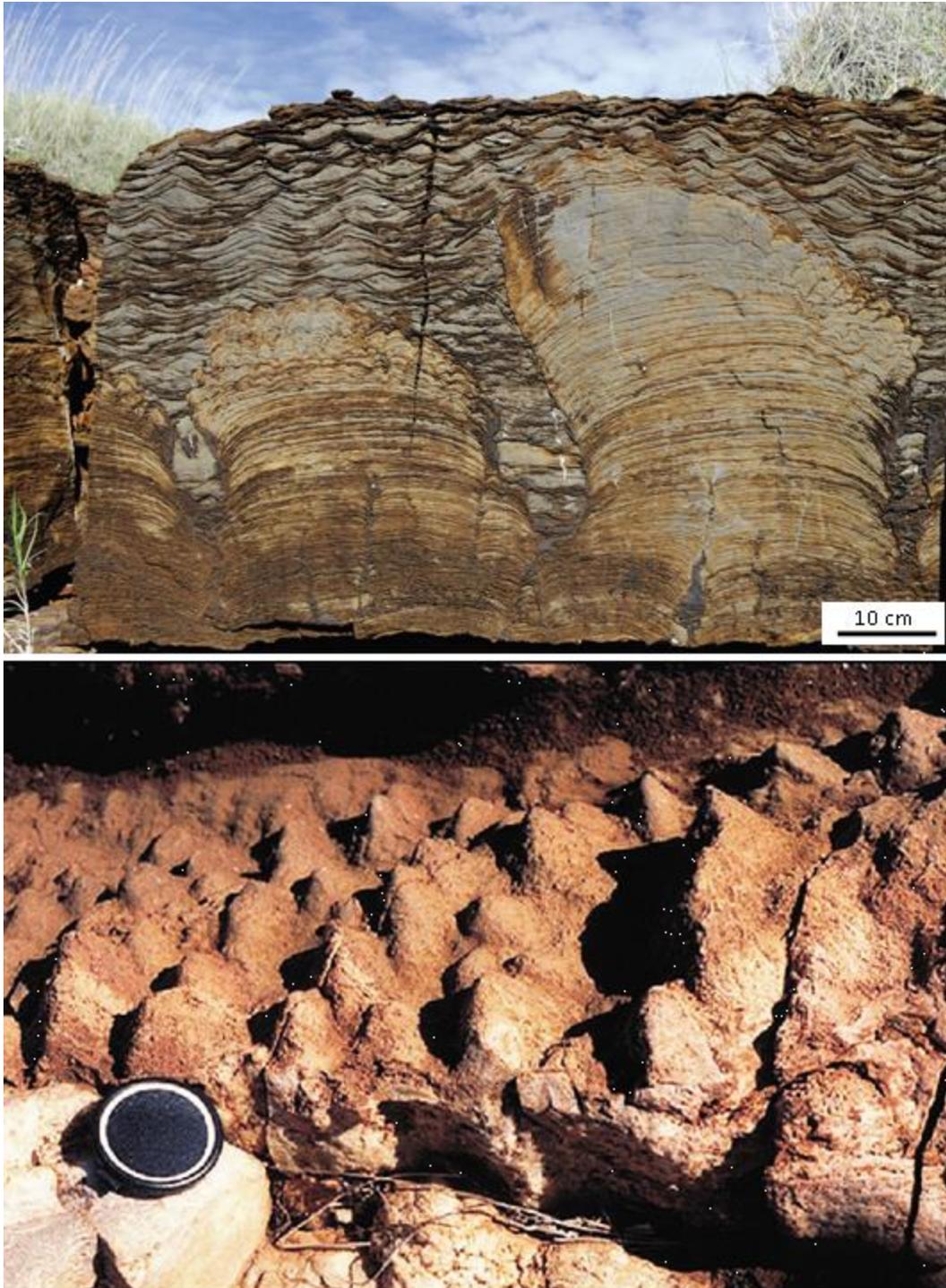


Figure 3: Stromatolites often grow in close proximity to one another. A) Large, broad columnar stromatolites from c. 2.7 Ga Meentheena Member, Tumbiana Formation, Western Australia. B) 'Egg-carton' stromatolites from c. 3.43 Ga Trendall Locality, Strelley Pool Formation (sample now housed at the Western Australian Museum); lens cap for scale (6 cm diameter). Photo credit: A, SM Awramik (in Grey & Awramik, 2020); B, Hickman et al. (2011).

Question 3. Are there abiotic sources for your detection?

There may be abiotic sources for detections at each stage of evaluation, from the initial detection in the field, to the search for supporting lines of evidence in the laboratory. For example, metamorphism of sedimentary rocks can mimic some macromorphological features of stromatolites (Allwood et al., 2018; Zawaski et al., 2020). It is necessary to ascertain the deformation history of the unit to assess the likelihood that the laminated structure formed via metamorphism. If the detected structure is in a highly deformed or metamorphosed rock, this might suggest structural deformation is responsible for its formation (Allwood et al., 2018; Zawaski et al., 2020, 2021, although see also: Nutman et al., 2019, 2021).

Abiotic sedimentary processes common to benthic environments may also produce structures that are morphologically similar to stromatolites (e.g., fractal geometry work by Verrecchia, 1996, and Grotzinger & Rothman, 1996; radiating crystal splays in Chafetz & Guidry, 1999; laboratory experiments by McLoughlin et al., 2008). The possibility of abiotic formation of laminae is especially relevant in environments where it may be harder to differentiate the influence of biology on structure formation (e.g., in an evaporative pool within the supratidal zone, or in sinter splash zones in a hot spring system). Careful morphologic, petrographic, and geochemical analyses, coupled with morphogenetic hypothesis building, can help differentiate between biotic and abiotic structures (e.g., Flügel & Munnecke, 2010; Grey & Awramik, 2020).

Question 4. Is it likely that life would produce this expression in this environment?

Context is critical in identifying whether a laminated structure may be a stromatolite. Firstly, assess the source of the detection in the context of the regional and local geology. As stromatolites are sedimentary structures, it is essential to verify that the structure formed within a sedimentary deposit. Stromatolites are known from a wide variety of sedimentary settings, including marine, hypersaline, lacustrine, fluvial, hot spring, and cave environments, as well as potentially in soils (Bosak et al., 2013; Grey & Awramik, 2020, and references therein). The detection does not have to be from one of these environments, however, the environment in question should have been habitable for a long enough time to enable the production and preservation of a biological signal.

Stromatolite morphology is strongly influenced by its environment. It is requisite to assess whether the morphology and/or size of the laminated structure varies, and whether this change is consistent with the physico-chemistry of the paleoenvironment and interpreted water depth (Allwood et al., 2006, 2009; Petroff et al., 2010; Flannery & Walter, 2011; Riding, 2011; Bosak et al., 2013; Grey & Awramik, 2020). By recognizing whether the study site is sedimentary, retains biotic and environmental signals, and is conducive to stromatolite growth, one can

determine how likely it is that stromatolites could have occurred in that particular environment.

Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?

Steps 2-4 outline essential lines of evidence in establishing whether a laminated structure is biological in origin (Fig. 4). Other analyses, although not essential, can be used to strengthen and provide confidence in a biogenicity claim. For example, the presence of syngenetic, biogenic organic matter, detected through Raman or other laboratory analyses, can support a biogenic interpretation (see section on organic matter worked example; Pasteris & Wopenka, 2003). Moreover, stable isotope data can support the biogenicity of organic matter and/or minerals within laminae (e.g., Flannery et al., 2016; Marin-Carbonne et al., 2018). The presence of microfossils within the laminae of stromatolites can further support a biological interpretation, although their preservation potential in ancient carbonates, in particular, is considered low (e.g. Buick et al., 1981).

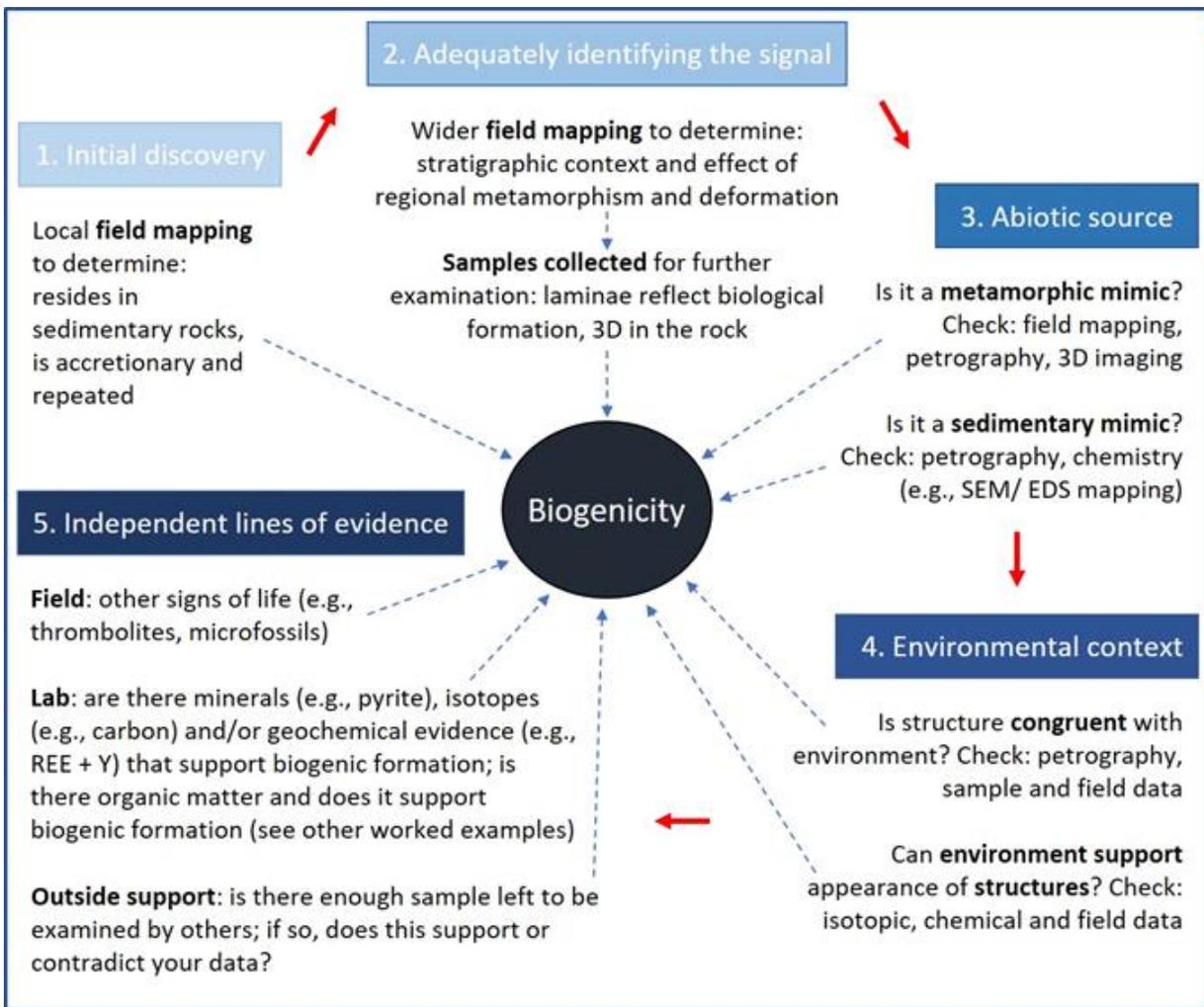


Figure 4: *Summary diagram highlighting some of the necessary actions (including suggested analyses) required at each step of the framework when using it to assess an in situ detection of an ancient laminated structure. In this context, it is important to recognize that there are multiple lines of evidence at each step in the framework and that each of these contribute to the biogenicity assessment. As you work through the framework, the relative confidence in a detection being biogenic increases (as indicated by the increasing intensity of blue in the headings).*

When looking at an individual line of evidence, there may be an abiotic explanation. However, it is unlikely that multiple lines of evidence supporting a biogenic origin will exist in the same sample in the absence of life (Fig. 4). Thus, multiple independent and complementary analyses can improve the confidence level of a detection and better assess the validity of a biological explanation.

3.2 In situ Worked Example: Detection and Characterization of Organic Deposits on Early Earth and Mars

Understanding the origin and alteration of organic matter on Mars as well as early Earth is of significant importance. Organic molecular structures have the capacity to record substantial amounts of information that could indicate their origins, whether abiotic or biological (see Fig. 5). Like early Earth, ancient Martian sedimentary environments exhibit the potential to have preserved ancient crustal organic deposits. A key current area of investigation in this area is selecting the most promising sedimentary Martian samples for collection by the Mars Perseverance rover for future return to state-of-the-art laboratories on Earth.

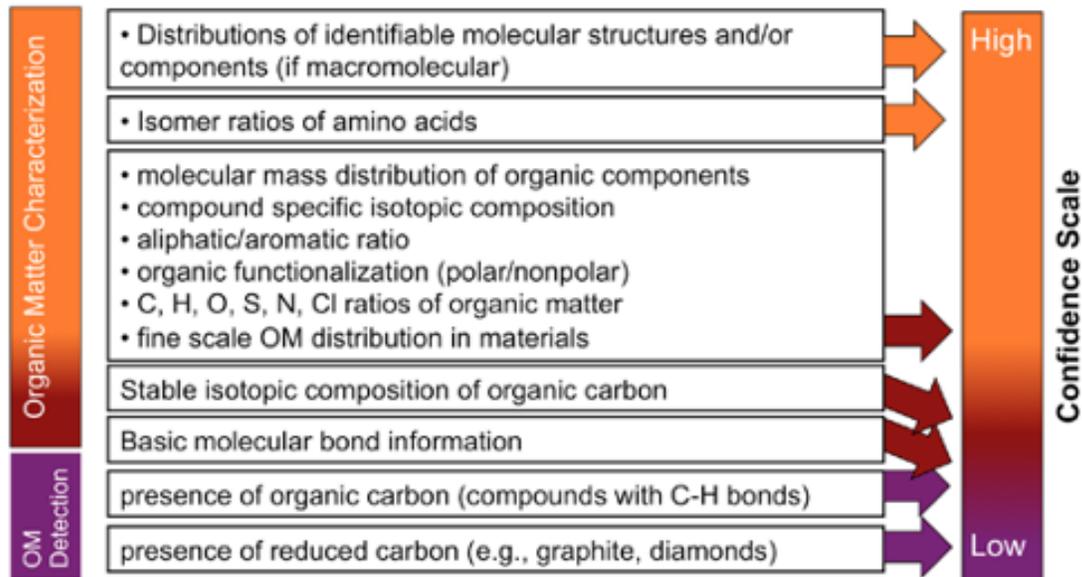


Figure 5. Organic molecule detection methods are not as definitive as some types of organic matter characterization. Confidence of detecting a definitive biosignature for various observation types for organic matter. Observation types (central column) in order of certainty that the observation could yield a definitive biosignature (right column). Left column denotes general type of observation: detection vs. characterization. (Adapted from Mustard et al., 2013).

Question 1. Have you detected an authentic signal?

The characterization and identification of organic compounds from geologic materials from both Earth and Mars generally rely on mass spectroscopic techniques that can be affected by both sample handling procedures and instrument parameters. For example, the potential for a false positive can arise when tuning is non-optimal resulting in fragmentation that is ambiguous or difficult to interpret. These erroneous signals differ from contamination since they represent either machine noise or degradation of components that can be mistaken for organic compounds of interest. Careful tuning and maintenance, both in the laboratory and on flight instruments, as well as procedural blanks and standard characterization, can prevent misidentification and indicate unreliable origins of a signal.

Question 2. Have you adequately identified the signal?

On Earth, sedimentary organic matter is the most abundant, widespread, and ancient remnant of our early biosphere. Macromolecular reduced carbon (i.e., 'kerogen') is the most abundant organic component and occurs in rocks as old as 3.46 Ga (Alleon et al., 2019). Solvent-extractable organic molecules that are demonstrably indigenous to their host rocks have been confirmed in deposits as old as 1.6 Ga (Summons et al., 1988). Thermal metamorphism presumably has obliterated these molecules in older deposits, possibly even converting such material to graphite (van Zuilen et al., 2002). Multiple laboratories have utilized

multiple techniques (e.g., GC-MS, laser raman and NMR spectroscopy, etc.) to corroborate detections of kerogens and solvent extractable hydrocarbons that occurred principally in shales spanning a range of thermal maturities and representing billions of years in age (e.g., Schopf and Klein, 1992).

On Mars, the Sample Analysis at Mars instrument (SAM) on Curiosity rover analyzed multiple portions of the drilled Cumberland mudstone sample in Gale crater and detected several chlorinated hydrocarbons, including chlorobenzene and two- to four-carbon dichloroalkanes, at tens of parts-per-billion (Freissinet et al. 2015). Other ancient Gale crater mudstones yielded a range of other organic compound classes, including thiophenic, aromatic, and aliphatic compounds released from the samples at elevated temperatures >500C (Eigenbrode et al. 2018). Through multiple procedural blank runs and supporting laboratory analog experiments, the SAM team identified an array of contaminants sourced from the instrument, including several compounds resulting from reactions between the samples and a SAM chemical derivatization agent (MTBSTFA) and successfully distinguished them from organic compounds that were indigenous to Mars. Laboratory simulations indicated that the chlorobenzene detected by SAM was derived from reactions during pyrolysis between perchlorate salts and martian benzoic acid in the mudstone (Freissinet et al. 2020). The sources of the organic carbon (i.e., compounds with C-H bonds) are currently unknown and could derive from a variety of Martian (igneous, hydrothermal, atmospheric, or biological) or exogenous sources such as carbon-rich meteorites or interplanetary dust particles.

Question 3. Are there abiotic sources for your detection?

On Earth, reduced carbon constituents in sedimentary rocks are derived ultimately from biological sources; however, some specific kerogen macromolecular structures and solvent-extractable compounds can be created by the chemical alteration and degradation of these constituents by thermal metamorphism, an abiotic process (Sephton & Hazen, 2013). Examples of demonstrably abiotic synthesis of organic compounds are sparse and perhaps limited to low-molecular weight alkanes in the deep continental crust (e.g., Sherwood Lollar et al., 2002) and hydrothermal systems (e.g., Welhan & Craig, 1982; Ménez et al., 2018). Abiotic organic material of meteoritic origin has also been identified concentrated within a specific horizon of the 3.33 Ga Josefsdal Chert (Gourier et al., 2019).

On Mars, the organic molecules in the samples analyzed by SAM could have derived from a variety of abiotic sources of either Martian or exogenous origins. It is difficult to assess the likelihood that the organics detected by SAM are abiotic or biotic in origin without additional characterization of the organic molecules, which is not possible with the current instrument suite. Accordingly, the SAM detections of organic molecules on Mars will remain on the low end of the biosignature confidence scale.

Question 4. Is it likely that life would produce this expression in this environment?

On Earth, sedimentary rocks that harbor the greatest abundances of organic matter were deposited in environments that favored both relatively high rates of primary production and enhanced deposition and preservation of organic matter in clay-rich sediments (e.g., coastal environments, estuaries) (Hedges & Keil, 1995). Given the highly efficient remineralization of organic matter by microorganisms, the abundances of organic matter in sedimentary rocks are largely determined by the efficiency of their preservation in sediments.

On Mars, the fine-grained sedimentary rocks (mudstones) analyzed by Curiosity in Yellowknife Bay, Gale crater, were formed in an ancient lake environment that would have been suited to support a Martian biosphere (Grotzinger et al., 2014). The aqueous environment was characterized by neutral pH, low salinity, and variable redox states of both iron and sulfur species. Carbon, hydrogen, oxygen, sulfur, nitrogen, and phosphorus were measured directly as key biogenic elements, alongside a variety of organic molecules detected in the ancient mudstones. Similar lake environments might have been relatively widespread on early Mars.

Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?

On Earth, independent lines of evidence include co-occurring biological structures, biogenic minerals, inorganic metabolites, and stable isotopic patterns that collectively increase confidence that accompanying organic deposits have biological sources (see Fig. 5). Multiple laboratories utilize arrays of techniques (e.g., GC-MS, Raman and NMR spectroscopy, etc.) to characterize demonstrably biogenic molecular structures (e.g., hopanes, steranes, porphyrins, carotenoids, etc.) situated in solvent-extractable phases and covalently bound in kerogens.

On Mars, more information is needed about the composition and structure of the organic matter detected in the mudstones prior to pyrolysis heating, which can lead to significant modification and destruction of the organics. Chemical analyses of the full distribution of solvent extractable organic molecules, compound specific stable C, H, O, and S isotope measurements of organic matter, and spatially resolved measurements that could reveal organic mineral associations in the samples would all be important in efforts to constrain their origins (abiotic or biotic). These measurements are beyond the capability of the Curiosity payload. Given the SAM instrument's limitations and additional constraints associated with in situ missions (limited time, power, and resources to analyze numerous samples and repeat measurements), the unambiguous detection of an organic biosignature will pose great challenges in situ, and may be impossible. The search for definitive

evidence of chemical biosignatures is one of the strongest motivations for Mars Sample Return (Meyer et al., 2021).

3.3 Remote Sensing Worked Example: Oxygen Detection on Exoplanets

Diatomic oxygen (O_2) is considered a compelling target biosignature molecule in exoplanetary atmospheres because of its strong fundamental absorption bands, clear link to biological production, large atmospheric and geologic sinks, and comparatively small abiotic sources (Owen, 1980; Sagan et al., 1993; Des Marais et al., 2002). However, recent research has identified several possible false positive mechanisms for abiotic production of oxygen on exoplanets, especially those orbiting M dwarf stars (Domagal-Goldman et al., 2014; Tian et al., 2014; Wordsworth & Pierrehumbert, 2014; Harman et al., 2015; Luger & Barnes, 2015; Hu et al., 2020).

Here we consider the application of the biosignature Assessment Framework for the 0.76 μm O_2 -A band on a terrestrial planet orbiting in the HZ (Kopparapu et al., 2013; Kasting et al., 2014) of a Sun-like star. Fig. 6 shows an outline of the process following the Assessment Framework, and applicable actions are described in more detail under each of the five framework steps below.

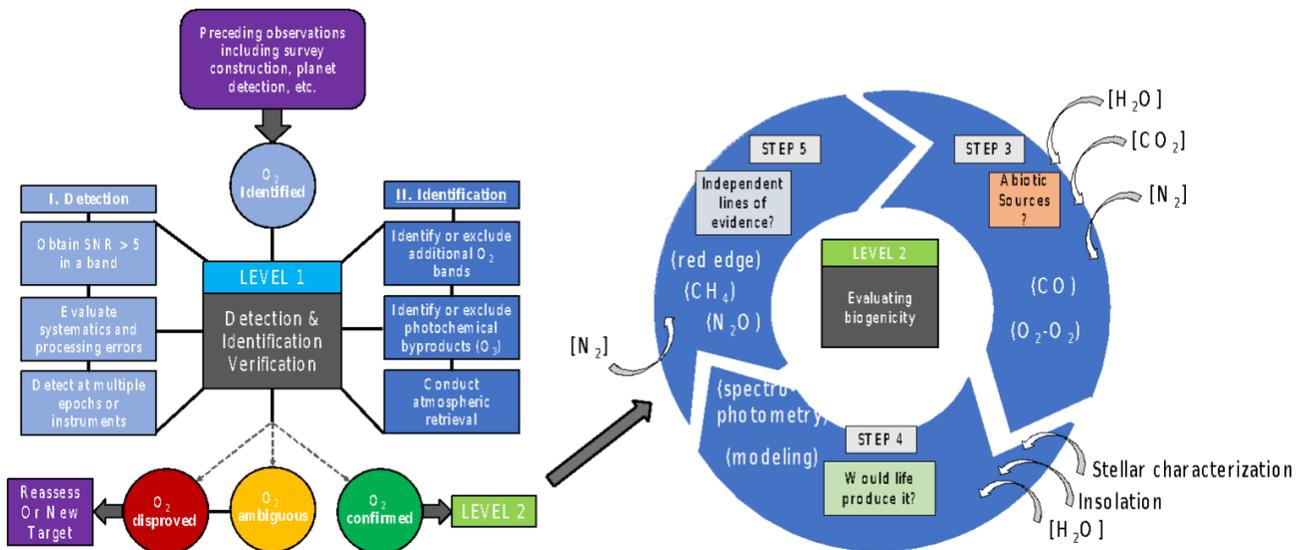


Figure 6: Schematic process diagram for assessing an initial detection of exoplanetary O_2 as evidence for a global oxygenic photosynthetic biosphere.

Question 1. Have you detected an authentic signal?

Validation of an O₂ detection requires obtaining a sufficient signal-to-noise (SNR) to rule out a spurious signal from an instrumental artifact. A SNR of 5 (Tessenyi et al., 2013) is considered a reasonable threshold for detection of a spectral feature, although a single band detection is more vulnerable to being confused with artifacts due to background, instrumental and processing effects, as well as stellar systematics. Ideally, potential artifacts should be ruled out via internal vetting and external independent data analysis. The initial interpretation can be strengthened by detections of multiple bands at different wavelengths, comparison of the suite of observations to models, and retrieved abundance estimates. The confidence in single or multiple band detections would also be enhanced by repeat detections over multiple epochs by the same observatory and/or instrument. Confirmation with a different instrument or telescope is likely not immediately feasible for a space-based detection, but could later be pursued to enhance confidence in the detection. Ground-based detections may be more immediately validated via independent observation. Improved modeling and spectral retrieval schemes will support more accurate detection and identification of potential biosignatures.

Question 2. Have you adequately identified the signal?

The O₂-A band absorbs in a region of the electromagnetic spectrum relatively free from other common absorbers, which helps to minimize misidentification (Des Marais et al., 2002; Schwieterman et al., 2018). However, detecting other O₂ and O₃ bands (see Table 2 in Meadows, 2017; Table 4 in Catling et al., 2018) at adequate SNR will increase confidence in the identification, as the characteristic pattern of O₂ absorption is unlikely to be caused by noise or instrumental systematics. Additional O₂ bands include the 1.27 um, 1.06 um (O₂- O₂ CIA), 0.69 um, and 0.63 um bands (Gordon et al., 2017). At very high O₂ abundances, O₂- O₂ CIA may also be seen at 0.445, 0.475, 0.53, 0.57, and 0.63 um (Richard et al., 2012; Karman et al., 2019), though this may be indicative of a false positive scenario (see Step 3). O₃ is an expected photochemical byproduct of O₂ photochemistry, and detection of ozone (O₃) in the UV (0.15-0.35 um), visible (0.45-0.85 um) and mid-infrared (9.6 um) would further support the inferred presence of O₂ in the atmosphere (Leger et al., 1993, 2011; Reinhard et al., 2017; Defrère et al., 2018; Olson et al., 2018; Quanz et al., 2018; Line et al., 2019). Potential contamination by telluric O₂ lines in ground-based high-resolution observations can be carefully removed via the planet's characteristic orbital velocity behavior (Snellen et al., 2013; Serindag & Snellen, 2019).

Question 3. Are there abiotic sources for your detection?

Abiotic sources may be distinguished by acquiring additional planetary and astrophysical contextual information. By systematically modeling the larger photo-

and geo-chemical environment, we can evaluate how likely it is that the oxygen is produced either by the photolysis of CO_2 (Hu et al., 2012, 2020; Tian et al., 2014; Harman et al., 2015) or the photolysis of water and subsequent hydrogen escape (Luger & Barnes, 2015; Wordsworth et al., 2018; Krissansen-Totton et al., 2021). These abiotic mechanisms have also occurred on Mars and Venus. On Mars, atmospheric recombination of photolyzed CO_2 is slowed by the lack of water vapor (Nair et al., 1994; Zahnle et al., 2008), allowing this abiotically-generated O_2 to build up to 0.1% of the atmosphere, a mechanism that may be even more pronounced on planets orbiting M dwarfs (Gao et al., 2015). Whereas on Venus, the early runaway greenhouse, water photolysis and H loss may have generated an O_2 -rich atmosphere (Chassefière et al., 2012) that was subsequently sequestered into a magma ocean or the crust, or lost to space (Gillmann et al., 2009; Luger & Barnes, 2015; Schaefer et al., 2016; Wordsworth et al., 2018).

Distinguishing between a true global biosphere and abiotic false positive scenarios may be possible with additional observations of the environmental context (Meadows, 2017; Meadows et al., 2018). These strategies extend to co-detecting additional biosignatures (Seager et al., 2012; Kaltenegger, 2017; Schwieterman et al., 2018) and establishing quantitative, Bayesian frameworks for remote biosignature assessment that help assess whether the signal is more or less likely to be due to life (Catling et al., 2018; Walker et al., 2018). Characterizing the host star's X-ray and ultraviolet radiation, which can vary with stellar age and type of the host star (Lloyd et al., 2016; Hidalgo et al., 2018) informs the likelihood of photochemical generation of abiotic O_2 . For a Sun-like star, many known false positive scenarios become less likely (Meadows, 2017; Meadows et al., 2018). Constraints on the abundance of O_2 from the presence or absence of O_2 - O_2 collision-induced absorption (Fig. 7; Schwieterman et al., 2016) would indicate whether the planet has a massive post-ocean-loss O_2 atmosphere. Quantifying atmospheric CO_2 , CO and H_2O , would help quantify limits on the possible photochemical production of O_2 (Fig 7; e.g., Schwieterman et al., 2016).

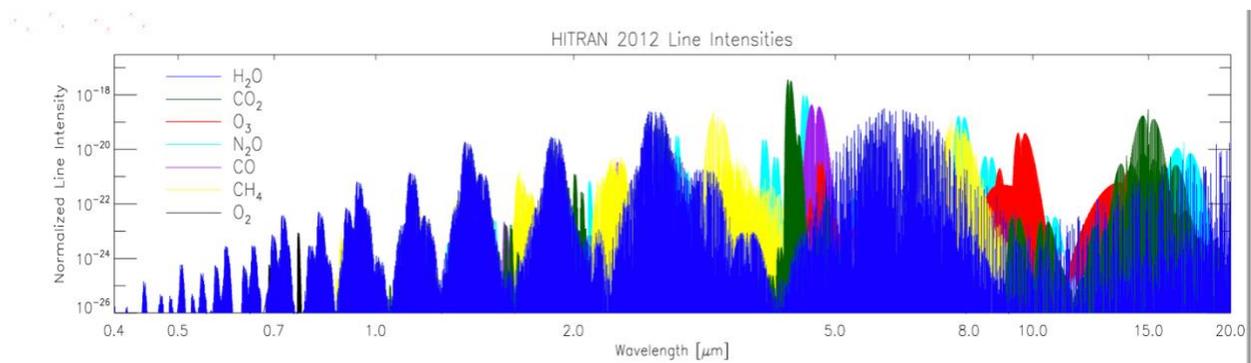


Figure 7: Combined spectral feature catalog for O_2 - O_2 , CO_2 , CO , H_2O and other molecules.

Question 4. Is it likely that life would produce this expression in this environment?

The direct characterization or inference of surface habitability would demonstrate congruence between a putative biosignature detection and the life interpretation (see Robinson & Reinhard, 2020, for a review). If the exoplanet with the potential O₂ biosignature is observed within the habitable zone, it will receive stellar insolation that increases the chances for a habitable climate, but follow-up measurements would be needed to determine whether the planet is, in fact, habitable. Direct detection of surface liquid water could be attempted via glint detection near crescent phase (Robinson et al., 2014), or a combination of surface-type mapping and phase-dependent changes in reflectivity (Lustig-Yaeger et al., 2018). Detection of water vapor coupled with surface temperature and pressure estimates from atmospheric retrievals could also be used to confirm habitability (Feng et al., 2018). The evaluation of biosignature congruence with a habitable environment is tied to the environmental assessment for false positives in Step 3 and may require substantial iteration.

Steps could also be taken to determine whether the observed abundance of O₂ or other gases are indicative of plausible biological fluxes or build up over time. The availability of metabolic substrates for oxygenic photosynthesis, such as H₂O and CO₂, could be quantified, and the history of atmospheric generation or loss of O₂ due to stellar evolution could be modeled (Luger & Barnes, 2015; Schaefer et al., 2016) and compared to the measured O₂ abundances. Whether biological oxygen accumulation is possible given plausible organic burial rates, oxygen sinks, and photon availability could be modeled (Catling & Claire, 2005; Lehmer et al., 2018). If corroborating evidence for life is found (Assessment Step 5; e.g., coexisting atmospheric CH₄), the abundances of the other biosignature gases could be inferred and compared to plausible biological fluxes.

Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?

Among the most compelling independent lines of evidence to support a biological explanation for exoplanetary O₂ would be co-detected biosignatures, including atmospheric disequilibrium (Lovelock, 1975; Sagan et al., 1993; Krissansen-Totton et al., 2016, 2018) through co-detected CH₄ or N₂O, surface photosynthetic pigments, and seasonal variations in biosignature gases. The O₂/CH₄ disequilibrium would rule out many false positive scenarios (Step 3) due to the required replenishment fluxes of both species, which varies due to different photochemical destruction rates from different host stars. The 1.65 μm CH₄ band would be an important target for space-based reflected light spectroscopy, although detecting analog Proterozoic and modern Earth CH₄ abundances (< 1.8 ppmv) will likely be challenging (Reinhard et al., 2017; Kawashima & Rugheimer, 2019), but more favorable for Earth-like planets orbiting K dwarf hosts (Arney et al., 2019). N₂O (7.8 and 8.6 μm) would provide additional evidence for a surface biosphere (Rauer et al., 2011) and may be observed with a space-based

interferometer (Quanz et al., 2018) or in transit for an appropriate target (Meixner et al., 2019). N₂- O₂ dominated atmospheres (Stüeken et al., 2016; Lammer et al., 2019; Sproß et al., 2021) especially in the presence of surface liquid water (Krissansen-Totton et al., 2018) may be indicative of biological replenishment. The N₂ may be inferred from its photochemical products such as NO and HNO₃ (Tabataba-Vakili et al., 2016) or its 4.15 μm N₂- N₂ CIA feature (Schwieterman et al., 2015). A surface biosphere may be directly detected via surface pigment signatures such as the vegetation red edge (VRE) or analogs thereof (Seager et al., 2005; Kiang et al., 2007; Lehmer et al., 2021) or alternative pigments and reflectance features (Hegde et al., 2015; Schwieterman et al., 2015). Because pigment signatures occur in the same wavelength regime as major O₂ bands (see Assessment Step 1), they can be searched for in contemporaneous UV-VIS-NIR reflected light observations, although VRE-like signatures will likely be significantly harder to detect than the O₂-A band (Brandt & Spiegel, 2014). Seasonal variation in biosignature gases may also provide additional confirmation of the biogenic nature of the gas (Olson et al., 2018), though quantitatively measuring seasonal changes may be implausible for most scenarios. Ultimately, obtaining independent lines of evidence for the biogenicity of an O₂ detection will be highly dependent on the available observing platform(s), as the time between follow-up missions will likely be long. Nonetheless, an O₂ detection and confirmation that passes through Assessment Steps 1 through 4 would warrant the additional resources required to search for corroborating signatures. Lastly, biological sources other than photosynthesis (see above) will need to be considered.

3.4 An Agnostic Biosignature Worked Example: Distribution of Enantiomeric and Isotopic Composition of Chiral Molecules

The search for evidence of extraterrestrial life in the Solar System has been guided by our understanding of biological processes and, in particular, by direct analogy to specific observations of the chemical and physical expressions of life on Earth. This strategy certainly has merit when exploring nearby terrestrial planets such as Mars, that may share biochemical heritage with life on Earth due to exchange of surface material over billions of years, but becomes less practical when searching for life in the outer Solar System (e.g., Ocean Worlds) where common origins are less likely. The recent Astrobiology Strategy document (NASEM, 2018) specifically recommended that researchers expand efforts to identify and validate biosignatures that are agnostic to life's chemical or molecular framework, particular metabolisms or evolutionary endpoint (Conrad & Nealson, 2001; Cabrol et al., 2016; Johnson et al., 2018; Marshall et al., 2021).

The development of an agnostic biosignature system is a probabilistic process that considers the likelihood that certain chemical and physical structures would arise by abiotic processes by deriving first principles expectations for a system that relies on energy input unlikely without life. This requires embedded fundamental

mechanisms in an interpretive framework rather than analogy and/or anomaly detection. There is an inherent tension in this process, however, since the potential for false positives is greater when the expected signals are too general. Further, an agnostic biosignature framework requires an exquisite prior knowledge of abiotic processes, an area too often overlooked when establishing search parameters. Agnostic biosignatures build from established abiotic baselines and a knowledge of what expressions are not unique to life. This can include potential prebiotic chemistry while assuming nothing about shared heritage. A consistent finding discussed during the Standards of Evidence Workshop pertains directly to agnostic biosignatures - the need for a much better systematic understanding of chemical and geologic expressions in the absence of life (Chan et al., 2019).

A chemical biosignature system with distinct potential as an agnostic concept was described in Glavin et al. (2020). The authors propose a combination of chiral asymmetry, isotopic composition, and molecular distribution as a way to diagnose the biogenicity of a suite of molecules that could be applied to amino acids or polyols (Fig. 8). This system builds on the recognition of the overwhelming homochirality expressed in the molecular structures produced by life on Earth. This concept also takes into consideration the fact that racemic mixtures or chiral molecules with much smaller enantiomeric excesses alone may not be a diagnostic feature of life, since similar enrichments due to non-biological processes have been found in meteorites (Glavin et al., 2020). However, if enantiomeric excesses are found in an extraterrestrial sample, a contextual analysis of the isotopic composition of these compounds could be used to indicate if biological fractionation resulting in the isotopic depletion typical of enzymatic reactions has also occurred. A final piece of this system considers the distribution of structural isomers. Biological processes on Earth result in the production of a small number of isomers relative to the enormous diversity of possible compounds of any class, thus a limited distribution could confirm the biogenicity of a suite of chiral molecules. This worked example considers the imagined evaluation of an assemblage of amino acids derived from a geologic sample.

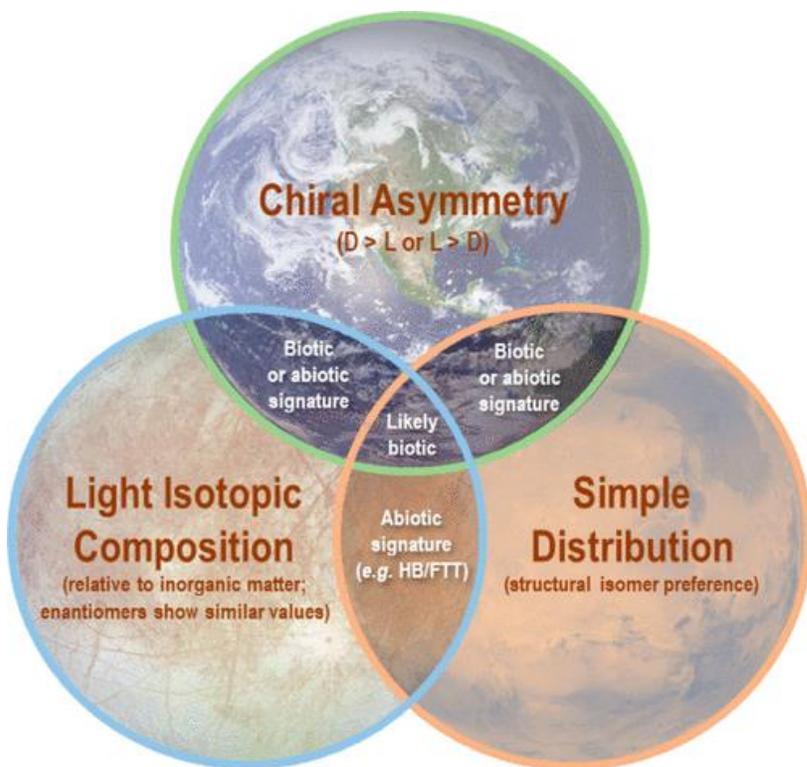


Figure 8. Venn diagram showing the three key molecular attributes (chiral asymmetry, simple distribution with structural isomer preference, and a light isotopic composition) needed to establish a likely biotic signature. Images adapted from Glavin et al. (2020).

Question 1. Have you detected an authentic signal?

The techniques necessary to measure both the stereochemistry and isotopic composition of a molecule, and to quantify the relative abundance of molecules, largely relies on different analytical methods, in particular gas and liquid chromatography mass spectrometry (GCMS and LCMS). Both methods also require solvent extraction of the molecules from the sample and chemical derivatization prior to analysis, types of mass spectrometry, etc. As with the early organic example, these methods require a well-maintained and tuned instrument calibrated with standards, a thorough characterization of instrument background, and replicate measurements in order to increase signal to noise and reduce uncertainty associated with molecular identification, enantiomeric ratios and isotopic composition. Compound specific stable isotopic measurements also require careful standard selection and data correction to avoid misleading fractionation values, since these measurements are reported in reference to an accepted standard (Hayes et al., 1989; Summons et al., 2006; Elsila et al., 2012). An authentic signal could be defined as a measurement above the instrument background (typically above a signal to noise ratio of ~3) and would need to meet precision and accuracy metrics that address background ion currents and shot noise (Merritt & Hayes, 1994).

Question 2. Have you adequately identified the signal?

In order for this system of stereochemistry, isotopic composition and abundance distributions to yield a conclusion with respect to biogenicity, the separation of enantiomers from other potential co-eluting compounds and their identification will need to be highly robust. In the case of amino acids, 96 amino acids have been identified by name in meteorites (including 12 of the 20 most common amino acids found in biology), although hundreds of different amino acids exist in these samples, but have yet to be identified due to a lack of analytical sensitivity and insufficient chromatographic separation for higher molecular weight species (Glavin et al., 2020). A general problem with GCMS and LCMS techniques is also the lack of commercially available standards that are needed to establish the identification of the unexpected or novel compounds we would want to include in an agnostic biosignature framework, particularly higher molecular weight compounds (Glavin et al., 2010; Cooper et al., 2016). If these short-comings can be statistically evaluated, the systematic differences between the racemic, undepleted, diverse assemblages typical of abiotic organic reactions will be distinctive from biological molecule assemblages that reflect evolutionary selection for constrained chemical properties that perform metabolic functions.

Question 3. Are there abiotic sources for your detection?

Agnostic biosignatures are the best demonstration of our need for better understanding of abiotic chemical and physical processes. In the case of this example, we can draw on the extensive characterization efforts that have gone into understanding the abiotic organic chemistry of meteorites, as well as recent research describing abiotic organic chemistry in subsurface and submarine hydrothermal systems. When diagnosing the biogenicity of a molecular assemblage by enantiomeric excess, it is important to remember that significant amino acid (and sugar acid) enantiomeric excesses of abiotic origin have been observed in extraterrestrial materials for some compounds (up to 60% L-amino acid excesses and 100% D-sugar acid enrichments). In addition, racemization of compounds of either biological or abiotic origins can occur over time due to exposure to liquid water and elevated pressures and temperatures during diagenesis that can completely erase any initial chiral asymmetry (Schroeder et al., 1998; see Fig. 6 from Glavin et al., 2020).

An illustration of several hypothetical relative abundance of enantiomeric proportions in an extraterrestrial sample, and possible origins, is shown in Figure 9. Since large L-amino acid excesses of abiotic origin have been found for a few protein and non-protein amino acids in meteorites, the simple detection of a large L-excess in a single amino acid may be insufficient to establish a biological origin. In contrast, the detection of a simple distribution of multiple amino acids that all

possess large L-excesses or any amino acids that have large D-excesses would provide more robust evidence for biochemistry, since most amino acids of abiotic origin are racemic, and to date, no significant D-amino acid excesses have been identified in an meteorite or extraterrestrial sample (Glavin et al., 2020). The detection of a D-amino acid excess of any magnitude would therefore provide strong evidence of extraterrestrial origin (since terrestrial biological L-amino acid contamination could be ruled out), and more significantly, of a potential second genesis of life in our solar system. The most challenging scenario would be the detection of a racemic mixture of amino acids as these compounds could be abiotic in origin or the result of complete racemization of an extinct biota. In this case, isotopic measurements could be critical to confirm a possible biological signal, but this requires an understanding of the isotopic composition of source reservoirs and not just a comparison to typical Terran metabolic values. Since racemization does not lead to measurable isotopic fractionation of amino acid enantiomers and chiral amino acids of abiotic origin in meteorites are all enriched in ^{13}C , ^{15}N and D relative to biological amino acids (Glavin et al. 2020), the detection of a highly depleted isotopic values in both enantiomers of a racemic amino acid mixture could provide strong evidence of an extinct biological source with the caveats mentioned previously. The best correlation to support a biological origin for an assemblage of amino acids would be a simple distribution of chiral amino acids with large enantiomeric excesses (either L or D) and with highly depleted isotopic values relative to inorganic matter in the sample. However, like enantiomeric excess, simple distributions of straight-chained amino acids with light carbon isotopic composition of these compounds can then be used to confirm a possible biological signal but requires an understanding of the isotopic composition of source reservoirs and not just a comparison to typical Terran metabolic values. The best correlation to support a biological explanation for an assemblage would be a depleted isotopic value of the more abundant enantiomer. Like enantiomeric excess, simple distributions can also be produced by mineral catalyzed or hydrothermal reactions (Islam et al., 2001; Burton et al., 2012); however, these conditions generally produce achiral products. Hence, the most reliable interpretations will rely on the combination of all three measurements.

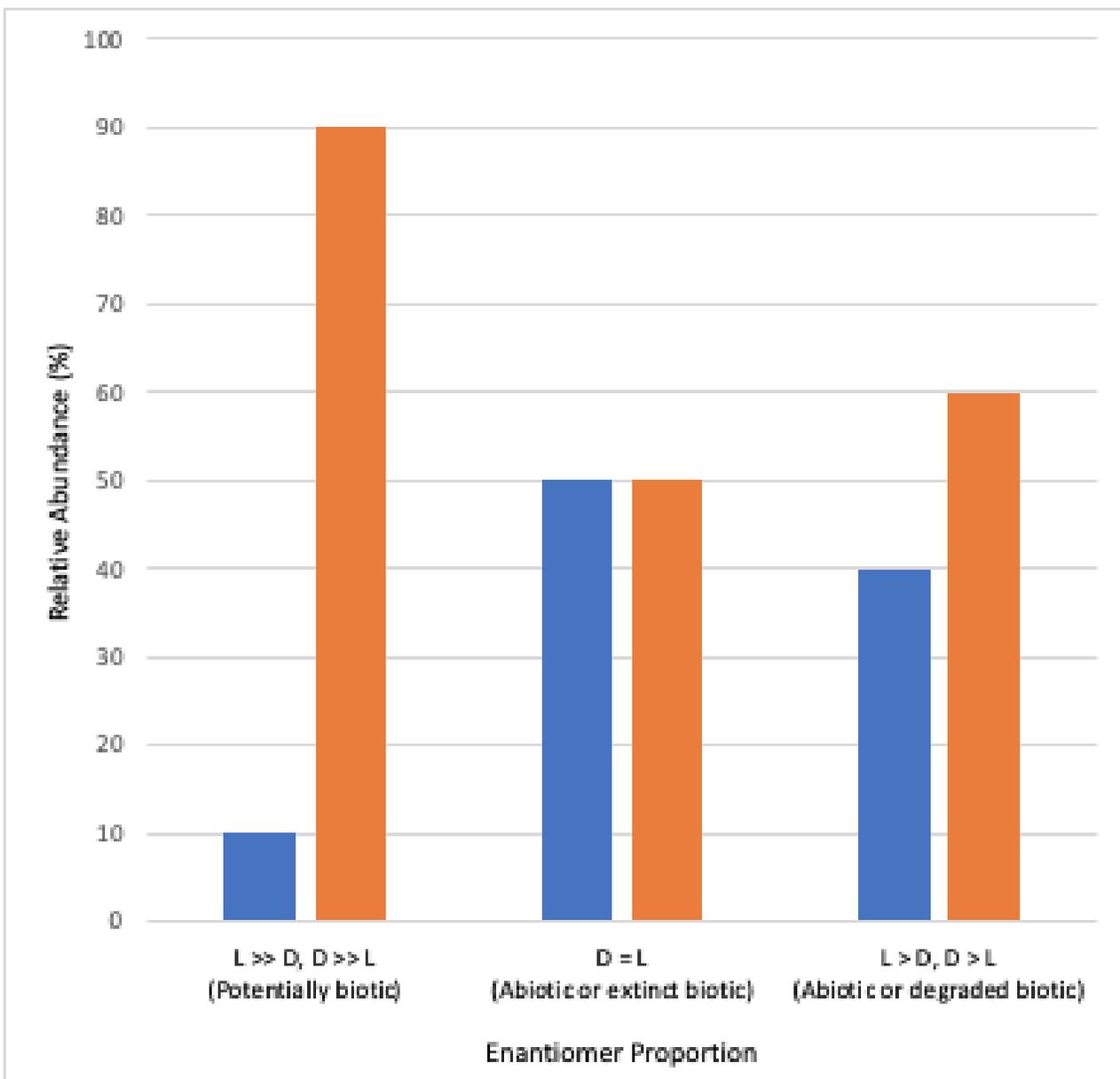


Figure 9. Some hypothetical relative abundances of D- and L-amino acid enantiomers and their possible biotic or abiotic origins (Adapted from Glavin et al., 2020).

Question 4. Is it likely that life would produce this expression in this environment?

Diagnosing the likelihood that compounds in a molecular assemblage would derive from a living system could be accomplished by considering the thermodynamic probability that the physico-chemical formation conditions would produce the same structures with similar stereochemical features. This observation would not require prior knowledge of the metabolism or chemistry of potential biota. Further, knowledge of the physical conditions could indicate if preservation of these

molecules is even likely or if an exogenous source is needed to achieve the observed molecular abundances.

Question 5. Are there independent lines of evidence to support a biological (or non-biological) explanation?

Given the numerous abiotic reactions that could produce these patterns, confirmation should include spatially resolved chemical, mineralogical, and isotopic studies to establish the origin of organic assemblages. Incubations or analog experiments could be performed to understand novel formation mechanisms or to see if there are physical parameters that promote production and can indicate either a chemical or a biological pathway.

4. Applications of the Framework to Missions

Section 3 provided case study applications of the Biosignature Assessment Framework for specific biosignatures and measurement techniques, in the scenario where a claim of a biosignature detection had been made. Here we take a generalized approach to how the Framework may be applied in a different way, to help guide the different stages of mission development, from concept to operation.

4.1 Utility of the Framework

Implementation of the Assessment Framework in mission designs starts with defining the science objectives, investigative strategy, concept of operations, requirements flow, and hardware build, followed by mission science return reporting and iterative data interpretation. Starting from conceptualization through operations and decommissioning; this Framework will be useful throughout the lifecycle of biosignature detection missions. Observatories that are not designed for life detection may benefit from this Framework to initiate such a search. Mitigating sources of contamination and establishing biogenicity are of paramount importance and the Framework can be a powerful tool. The Framework can also be used to anticipate experimental, methodological, and statistical needs that build in estimates of uncertainty for observations and interpretations.

4.2 Mission Sensitive Application of the Framework

Confirmation of a detected signal (Question 1) begins with understanding calibration, background noise, sensitivities, resolution, operations, and behavior of instruments under unforgiving environmental pressures of space. Establishing confidence in the signal (Question 2) requires a solid understanding of sources that could impact data interpretation. Pre-flight analog testing and stringent contamination control and knowledge procedures will help distinguish data artifacts and false positives. Confirmation of a detection during a mission can also be supported by multiplicate measurements with multiple instruments investigating the same target body. The ubiquitous presence of Earth's biosphere has limited our knowledge of the natural expression of abiotic physical and chemical systems,

however, false positives (Question 3) can be recognized by understanding measurement capabilities for known abiotic signals in a given environment. Question 4 requires that the Science Traceability Matrix (STM) connects measurements back to key aspects of habitability and the context of the planetary system in order to determine if the biosignature in question would likely be produced in this system. The STM should also assess sample concentration and preservation. Similarly, Question 5 requires that the renewed STM trace multiple measurement techniques to determine if independent lines of evidence support biological (or abiotic) formation. The value of independent lines of evidence, with detectable interdependence, consistently pointing in the same direction strengthens confidence in the overall conclusion.

4.3 Mission Life Cycle Considerations

Implementation of the Assessment Framework could prove useful throughout the lifecycle of missions, from conceptualization through operations and decommissioning.

During pre-Phase A, a mission concept is developed, often starting with a team that studies the mission concept, with the objective of understanding how difficult it would be to carry out the mission's science goals. The output from those teams outlines the science goals, defines options for mission architectures and instruments that could realize those goals, and prioritizes the technologies that would be needed to enable or enhance those architectures and instruments. NASA-led studies of this type are usually referred to as a "Science Definition Team" for planetary missions and as a "Science and Technology Definition Team" for astrophysics missions. During this phase of a project, the Framework would help teams determine which, if any, levels of the Framework would be part of the mission concept's goals and allow the teams' guidance on what specific questions and investigations would be required to meet those goals. The Framework could also help formalize and motivate decisions made during these studies.

As an example, consider the final reports from the LUVOIR and HabEx mission concept teams. Although the Framework did not exist when those studies were conducted, both teams included a search for biosignatures and studied concepts that would meet at least the first four stages of the Framework. Both teams studied architectures with wavelength ranges that would allow for the detection of biosignatures (Step 1) evident either on modern-day Earth (e.g., O_2/O_3) or thought to have been evident in earlier eons of Earth's history (e.g., CH_4 or an organic haze). They also explicitly ensured "contamination" (Step 2) would be limited by studying the potential for background objects to confuse their detections and by considering observing strategies that would eliminate the potential for light from the Earth or Moon to enter the instrument array. Both teams considered abiotic production mechanisms (Step 3) for the biosignatures they were designed to detect and interpretive strategies to identify these abiotic mechanisms. These strategies included both a wavelength range on the "biosignature detection" instruments that would add environmental context to the detection, and

requirements for a second instrument that would measure the UV flux from the star, which is an important energy source for abiotic production of atmospheric biosignatures. These studies made sure the instruments responsible for detecting these molecules had the required spatial resolution. Similarly, both missions included a separate set of observations to confirm the presence of liquid water on targets. With its larger aperture, LUVOIR is also capable of searching for biosignatures associated with pigments or surface vegetation and could begin addressing Question 4 of this Framework. LUVOIR and HabEx observations could complement one another (Question 5) if an ideal target becomes available. All of this - the detection of the biogenic molecules, avoidance of contaminant light sources, discrimination of false positives, and confirmation of global-scale surface habitability - were explicitly included in the teams' Science Traceability Matrices, at the level of top-level science goals that flowed down to specific instrument properties.

These instrument properties were in turn used to help determine the technologies that would enable or enhance the missions and ensure they could collect the requisite data. These technologies were subsequently available for funding through NASA's competitive Technology Development for Exoplanet Missions (TDEM) call. The totality of this was presented in the final reports available to the community. Additionally, this allowed engineering teams - led out of NASA GSFC in the case of LUVOIR and out of JPL in the case of HabEx - to provide details on the architectures and instruments that would close on the top-level science goals. This enabled multiple outside groups to create "ballpark estimates" for the complexity and associated costs of these projects, so that these ambitious science goals and their associated costs could both be considered as part of the Astro2020 Decadal Survey.

In later phases of a mission, the Framework will be similarly useful. The Framework can help determine the necessary technologies and help prioritize the funding of those development activities. Similarly, top level mission architecture decisions can be made with specific Framework goals in mind, including incorporating Framework goals as primary mission goals with requirements for selected instrumentation to meet. Such requirements are used as benchmarks for performance for the mission and its instrumentation throughout the design, fabrication, assembly, integration, and testing phases. Placing specific Framework steps in the mission's top-level science goals is the best way for an interdisciplinary team of engineers and scientists to design a mission that closes on the life detection goals laid out and identified by a concept study team.

During operations, the relationship between the Framework and the mission comes back to the concept of confidence – a measure of certainty on what has (or has not) been detected. This would dramatically impact mission operations planning, the publications from the mission, and the communication of results from the mission.

For an example of how this would work, consider a rover on the surface of Mars. To achieve Questions 1 and 2 of the Framework, a rover would need to have the ability to detect a signal from life while ensuring decontamination steps, calibrations, and blank analyses - all of which are necessary to demonstrate the status of the instrument signal, noise, and operations, as well as ensure that there are no forward contaminants that might compromise analyses. To address Questions 3 and 4 of the Framework would likely require a coordinated use of instrumentation, and potentially follow-up missions. Confirmation of a detection could also be supported by other measurements of the same feature by the same instrument, other instruments on the mission, and/or instruments of other missions.

The specific approach will depend on the nature of the possible biosignatures being investigated on a mission, and the known likelihood that they may also be products of abiotic formation processes. That said, all biosignature investigations will benefit from contextual measurements that support characterizing both the sample, as it may help understand other instrument results, and details of the potentially habitable or inhabited environment. In applying Question 4 of the Framework, the mission should trace selected contextual measurements back to key aspects of habitability that may have the greatest value in assessing whether life would produce the biosignatures being investigated. Question 5 of the Framework applied to *in situ* missions requires that the STM traces a biogenicity investigation to multiple measurement techniques to determine if independent lines of evidence support biological or abiotic formation.

As on Earth, it is expected that extraterrestrial life would impart many signatures that are independently detectable but exist interdependently. Thus, the value of independent lines of evidence (e.g., lipid patterns, elemental distributions, and microbial mat structures) supporting a consistent interpretation instills much more confidence in the overall conclusion compared to any single line of evidence. Further, any detectable interdependency among the biosignatures may provide yet another orthogonal interpretative path to assess confidence (e.g., distribution of amino acid skewed to complex structures, enantiomeric excess, and peptide chains; or cellular structures and/or motility).

Application of the Framework to *in situ* biosignature-seeking missions should be continuous throughout the life-cycle of a mission, which may be decades. Following mission design, decisions for modifications and refinements to the build (e.g., instruments and contamination control) and implementation plans (e.g., resources for cruise and science operation activities, concept of operations) need to consider the impact of changes on how well the mission maintains the capability to address the Framework. The Framework provides high-level guidance that if implemented throughout the mission process (i.e., concept, design, build, launch, cruise, science operations, data analyses, and reporting) enables mission scientists to establish a knowledge base from which to assess and understand confidence assigned to biological (or abiotic) interpretations.

4.4 Confidence and Statistical Inference

Drawing on lessons learned from previous life detection claims that, on reevaluation, rested on faulty statistical reasoning; attendees at the Standards of Evidence workshop agreed that researchers should collaborate with statistics specialists when reporting statistical claims. Statistical methodology is critical in both the detection (Questions 1 and 2) and assessment (Questions 2, 3 and 5) levels of the Assessment Framework.

4.4.1 Biosignature Detection

No experimental result or claim is complete without an attendant measure of its uncertainty. This is particularly important for biosignatures since sample sizes are likely to be small, evidence marginal, and impact enormous. As such, care should be taken to follow standard methods to report uncertainty clearly and to increase the reproducibility of claims (Rein, 2019). All experimental results contain uncertainty resulting from resolution limitations and instrument and methodological systematic error, as well as random process uncertainty due to finite sample effects and instrument precision, and uncertainty in inductive and statistical inference methods.

Scientific conclusions are reached through deductive and inductive reasoning, with induction being required to make conclusions from empirical evidence. Given that biosignatures are empirical in nature and, by definition, induction cannot prove a conclusion but only lend strength to competing hypotheses, all biosignature detection reports will be uncertain. Methods exist to quantify the contribution of each of these sources of uncertainty, and to measure the interaction and propagation of these sources of uncertainty when combined by multiple-layers of analysis.

Though there are well established inductive inference rules, such as the methods of residues, that will undoubtedly be included in any report of biosignature detection, statistical inference is the most widely used framework for quantifying inductive uncertainty due to random processes. The uncertainty here is the result of the limited sample sizes available to experimentation causing the measured values to diverge from the true or population values. Statistical methods are powerful but require specialized techniques for understanding the assumptions being made by the methods precisely (i.e., the number of ties present in two populations of values is critical to interpretation of a Wilcoxon test). Statistical methods can only be used to quantify uncertainty due to random processes, they cannot be used to quantify uncertainty due to systematic processes, or even detect the presence of systematic error.

4.4.2 Biosignature Assessment

Depending on whether a Frequentist, Bayesian, or Likelihood approach is being used to test a hypothesis, uncertainty quantification will usually be in the form, respectively, of p-values and confidence intervals; posterior distributions and their

credibility intervals; or likelihoods. Researchers should be conscious of the pitfalls in generating p-values in particular, such as the large sample effect. P-values should be presented in conjunction with the attendant effect size (Sullivan and Feinn, 2012). Confidence and credibility intervals provide a more robust method of quantifying uncertainty (though they depend on the underlying statistical methods being properly applied). Common pitfalls include multiple hypothesis testing without p-value correction (the Bonferroni correction is the simplest and most commonly applied method, though not the only method available) and overfitting polynomials to data or fitting data to too few data points (i.e., the number of measurements should be less than the number of fitted parameters). Multiple-hypothesis correction guards against post-hoc hypothesis testing, in which hypotheses are selected on the basis of their statistical significance (also known as p-hacking), and HARKing (Hypothesizing After the Results are Known), in which hypotheses are generated for testing after the experiments have been conducted (Bishop, 1990).

An essential component dictating confidence is the Bayesian likelihood of a given signal being produced abiotically, $p(\text{signal} \mid \text{no life})$. In many cases this factor is computationally intractable, and highly vulnerable to model assumptions. Correspondingly, historically many abiotic false positives were unrecognized at the time of original publication (oxygenated atmosphere false positives, Meadows, 2017; abiotic microorganism-like features; McKay et al., 1996; isotopic and mineralogical signatures; Alleon et al., 2019; etc). Using the Bayesian framework helps to mitigate uncertainties in individual target interpretation by addressing population-level questions, such as the likelihood that a survey has found life on at least one target, and makes inferences more robust.

A Bayesian statistical methodology has the advantage that assumptions and prior-information can be encoded into a prior distribution, allowing quantification of uncertainty to be driven by observation. Frequentist methods are equally powerful, but prior beliefs are encoded in the choice of statistical method. For example, a researcher who assumes a non-parametric population distribution will likely choose a Wilcoxon test over the t-test. Bayesian methods would include that assumption in the prior distribution. Whatever the statistical methods employed, the assumptions being made should be clearly stated so that the scientific community can evaluate the merits of those assumptions. Statistical methods should be used to strengthen inductive arguments and Bayesian, Frequentist, and other schools of analysis should reach the same conclusions when properly applied - except when data is so limited that statistical power is severely limited (Albers et al., 2018). When combining p-values from multiple independent variables in a meta-analysis to provide support to a particular hypothesis, Fisher's combined probability test should be considered.

Systematic uncertainty results from consistent error in an instrument or methodology. Since systematic uncertainty cannot be detected through statistical methods, it is critical that the experimental design be such that it is robust to systematic error. However, the potential impact of systematic error can be crafted

in tandem with instrument design. Sensitivity analysis before and after experimentation can bound the impact of systematic error and guide experimental design to mitigate it. That is, analysis should be performed that relates the magnitude of systematic error in an instrument to the ability of the experiment and data analysis procedures to differentiate hypotheses. Unlike uncertainty due to random processes, systematic uncertainty can be eliminated entirely by careful experimental setup or by *post-facto* correction.

Two commonly used uncertainty propagation methods are worst-case and differential error quantification. Worst-case analysis measures the maximum error that could be accumulated at each step - from instrument resolution and precision error. These bounds are determined by instrument calibration tests through confidence or credibility intervals for a multi-stage analysis and report the maximum uncertainty that could have accumulated using the additive rules of uncertainty, with appropriate unit conversions. When performing linear regression on data that includes measurement uncertainty, the standard way to retain information about uncertainty in the best-fit slope is to calculate the ratio of the best fit slope to the slope given by the worst-fit, where the worst-fit line is bounded by the measurement error bars. Farrance & Frenkle (2012) provide a review including a summary of terminology to be used in reporting and discussing measurement uncertainty. Biosignature scientists should also be familiar with the “Guide to the Expression of Uncertainty in Measurement” detailed in ISO 15189. By communicating that uncertainty through the confidence or credibility intervals resulting from appropriate statistical tests a researcher can place their discovery on a spectrum of confidence and provide a measure of certainty.

5. Reporting Protocols for Life Detection Claims

Verifying, reporting, and communicating evidence in the search for life to decision-makers, the scientific community, and the general public, involves consideration of the different means by which each audience obtains, interprets, and assesses information. Given the importance and profile of life detection announcements, it is vitally important that the community facilitate and incentivize a “reporting protocol” that represents *best practices* that offer clear, correct, concise information on the nature of the detection and resists misinformation and sensationalization. The Biosignature Standards of Evidence Workshop draws on lessons learned from previous scientific efforts involving large communities and wide public interest. We suggest a strong program of verification be adopted not only as a means to encourage use of the Framework Assessment but as a mechanism for standardizing the reporting protocols used to communicate life detection claims both within the scientific community and to the public.

5.1 Ideal goals for community biosignature verification and reporting

Verifying results and interpretation

The ideal goals discussed are to ensure that both the scope and level of certainty in reported results are appropriately perceived by both the scientific community

and the public. A team reporting a life detection claim is ideally large enough to accommodate the diversity of discipline specialists necessary to provide a thorough validation of the methodology and results. While the focused research areas that address questions of astrobiological interest may be small, these projects represent a broad variety of methods, instruments, and specializations that may have independent experts outside of the community. Independent verification should be stressed as a high priority for high-profile results. Workshop participants agreed that a mechanism that could fund independent verification and enable the concurrent publication of both discovery and verification papers would benefit biosignature science and strengthen claims as an overt demonstration of Question 5 in the Assessment Framework. Mechanisms to facilitate the sharing of data and results prior to publication should be explored as an important component of the verification phase of a life detection announcement. Workshop discussions did recognize that there is an inherent tension between the speed of sharing and the rigor of peer review, as well as a clear danger that pre-print materials as they currently exist can be over-interpreted as final results, and suggest that new mechanisms may be necessary to facilitate this information dissemination. Community collaborative networks (e.g., NExSS and NFoLD) could serve as nodes to coordinate these activities and also to provide a forum to verify the fidelity of results, as an extra step before the peer review process.

Reporting to the scientific community

Discussants proposed a number of modifications to the typical publication timeline. In addition to a “two-factor” publishing process that couples discovery and verification, there was support for incentivizing a system where claims should be reported in a stepwise manner with the discrete confidence estimates included. This would allow the community to respond to individual interpretations rather than the nested assumptions and anticipations of an aggregate study. Claims should be reported with an explanation of uncertainty rather than an arbitrarily derived numerical score or descriptor. This prevents misunderstanding based on in-group language use. Claims should be carefully reported with contextual knowledge (Question 4) to communicate the scope of the discovery and avoid potential confusion with previous claims. High-profile missions could address these issues by formalizing built-in structures that protect the rigor of the science and promote constructive criticism by engaging the broader scientific community in constructing alternate data interpretations.

Reporting to the public

In keeping with a general theme that biosignature science should rely more heavily on discipline specialists, discussants found that life detection claims would benefit from more collaboration with science communication specialists. Examples given included institutional press officers and respected science journalists, but it was also stipulated that there should be an effort to communicate the nature and details of the Assessment Framework to these groups so that they too can verify public information and to discourage the communication of pre-print, pre-reviewed results. This, again, requires prior coordination, possibly with a network of trusted

science communicators. Workshop Participants agreed that claims communicated to the public should be clear that discoveries are usually at the beginning, not the end, of a lengthy scientific process that will likely require subsequent measurements or missions and may result in modifications of current interpretations. Specialists that can provide clear and visual descriptions of the discovery and verification of the results should especially include information provided to policy and decision-makers so that the nature of the discovery is clear.

5.2 Current obstacles to ideal reporting protocol

In the current academic climate, there is immense pressure to publish as many papers and as rapidly as possible. This is particularly true when high profile results are considered. Given that these activities are directly tied to measures of job performance, employment opportunities and/or grant success, an incentive program that encourages researchers to follow the protocol would require a community-wide change in how we assess productivity. Funding agencies are incredibly necessary partners in helping institutions re-imagine how to reward the more incremental nature that should be part of high-profile scientific endeavors. The proliferation of private investments into academic research activities may also affect the motivation to follow the protocol. Shareholder benefits may play a bigger role than scientific reasons to communicate high profile discoveries. In these cases, it may be beneficial if community organizations (e.g., NExSS, NfoLD, NOW, ExoPAG, MEPAG, OPAG) proactively initiate communication with the teams receiving private grants, and invite them to participate in the protocol measures and share insights on how to implement the Assessment Framework.

One of the most challenging elements of a lengthy verification process is how to maintain a discovery embargo when large groups of scientists are involved. Strict non-disclosure agreements within a discovery and verification collaboration would be required. Further, the concern of “getting scooped” is a powerful disincentive and encourages rapid publication without data sharing for independent verification (see Fig 5.1).

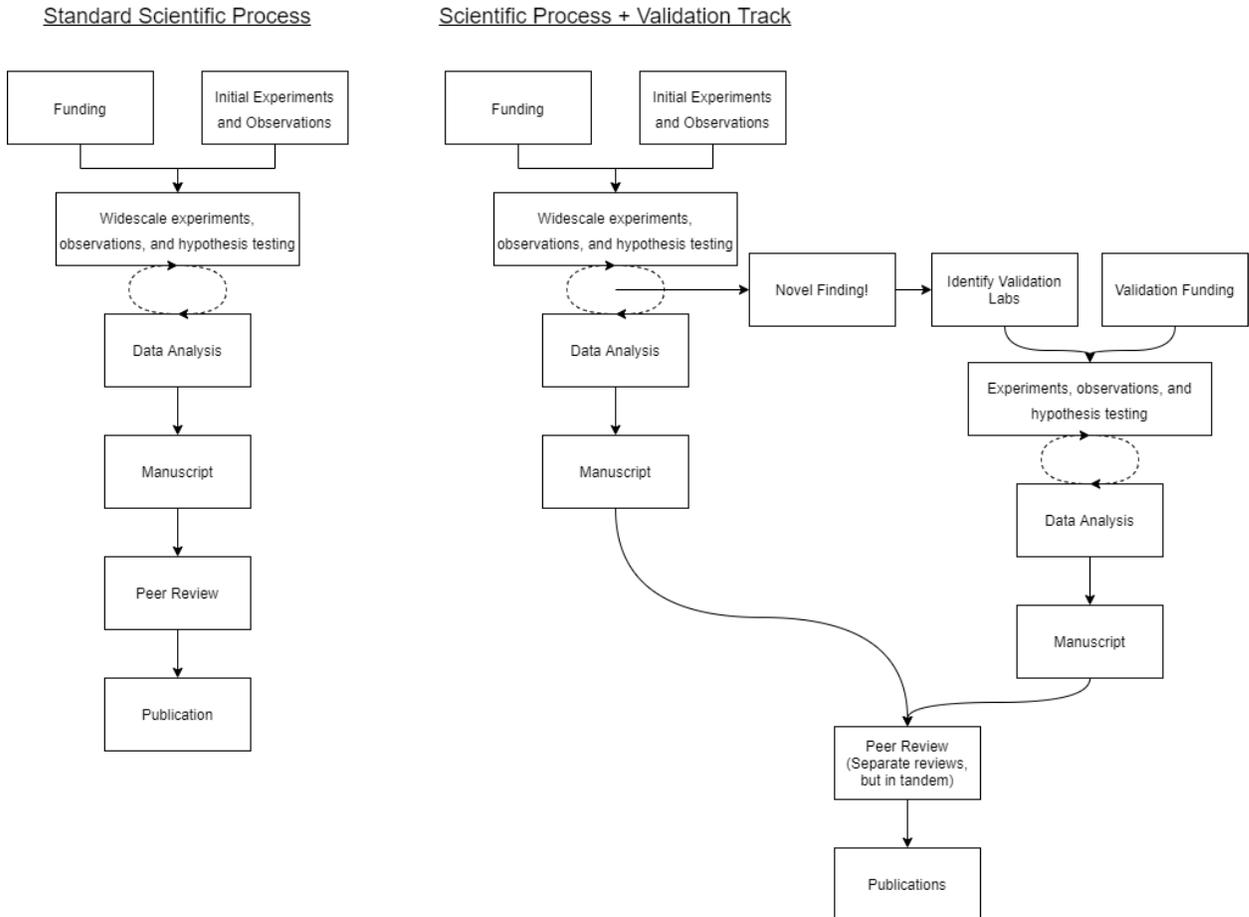


Figure 5.1. Proposed Workflow for Scientific Discovery Related to Life Detection Incorporating Outside Validation. In the left panel, the standard scientific process that leads to publication follows a predictable pathway. The right panel illustrates a new “validation track” whereby with initial pools of traditional funding, researchers will progress and iterate with increased data, leading to the testing of more precise hypotheses. If life detection is confirmed, funding agencies can then open a wider rapid search for more focused funding for interdisciplinary validation experiments.

Workshop participants agree that a community-led culture change is necessary to ensure that individuals are rewarded for their roles in a discovery in a way that also incentivizes a robust verification process. Individual institutions may also reward controversial results published in supposedly high impact journals to increase an institution’s profile rather than building incentives encouraging a robust scientific process. While communicating with the media for a wider exposure of the discovery results is well-intentioned, the pressure of high profile results and recognition beyond the scientific community, tied to career advancement, is a factor that could work against widespread adoption of the suggested verification protocol. Without strong community support there will be those that ignore the protocol for their own benefit. While there is no way to prevent this behavior, effective communication to researchers and institutions about the protocol's existence and benefits will be needed.

A final obstacle specific to large collaborations are the multiple, possibly conflicting sets of regulations by government agencies, national laboratories, and international collaborators. For example, it may be hard for U.S. civil servants such as NASA employees to participate in oversight activities. Additionally, all stakeholders would need to agree to the reporting protocol and fully understand how these processes would influence announcements.

5.3 Incentivizing the best possible verification and reporting process

For high profile science results, the community should foster an environment where the preference and priority is given to robust scientific processes. This would involve an inclusive collaboration built on trust that sharing data will not have detrimental outcomes, and an understanding of the common goal that a community consensus result will have a longer lifetime than a short-term sensational one that may be controversial. To achieve this, we may draw lessons from scholars who study ethics in astrobiology, social scientists, communications experts, and protocol specialists. The community may benefit from professionals who specialize in support and communication.

Incentivizing the verification process is a barrier to robust science because currently there is no direct reward or benefit for groups who verify results. A potential solution to this is a discretionary life verification detection fund — national agencies could have funds available for verification expenses, and acknowledge it as a funding source. The verifiers would be acknowledged in final results or perhaps in a companion verification publication. With proper funding it may be possible to build a verification process that does not significantly delay progress, and aims to reduce the time between claim and re-analysis while improving data and transparency. Engaging with the Assessment Framework and reporting protocols should be seen as a way to strengthen the legitimacy of bold claims and not as a regulatory hurdle that penalizes discovery.

Changes to the incentives for reporting life detection to the scientific community must be centralized at the journal level. Too often journals do not accommodate or publicize the caveats to a discovery with the same enthusiasm as an extraordinary claim. Workshop attendees also suggested a more broad adoption of the practice of publishing reviews along with papers as a way to broaden scientific discourse. Another mechanism discussed was a new type of publication that can serve as an interim report describing the progress an observation has made along the Assessment Framework and the steps that have been taken in an investigation. This would be another form of wider data sharing with minimal interpretation allowing broader introspection. Workshop participants also discussed the difficulty of reporting negative data or results that support a null hypothesis. Journal articles that can accommodate these reports or databases that can host null datasets are essential for helping researchers to recognize and evaluate potential experimental avenues and use this information to understand technological and biological limits.

Reporting high-profile results to the media and general public will require communication training that not all researchers are formally taught. Existing programs which do provide communications training include the Blue Marble Space Institute of Science Young Scientist Program and the Massachusetts Institute of Technology science journalism program. The community would benefit from astrobiologists engaging in these and others like them, with the incentive of gaining better communication skills and recognition. Further, workshop attendees were eager for more opportunities for researchers, science communicators and journal editors to discuss the reporting protocol and the perils of poorly verified publication that erodes public trust.

5.4 How do we encourage and support collaborative verification?

Community actions are the key to promote and support collaborative verification. Suggestions proposed during the workshop included:

- Dedicated sessions in scientific meetings that can be used to present the reporting protocol, discuss further improvements, establish verification mechanisms and demonstrate the benefits of following the protocol and participating in verification activities.
- Establishing a committee or group to handle implementation of the reporting protocol. The ideal committee should be international, sampling the demographics of our community and the diversity of disciplines involved in astrobiology. A useful step may be the formation of an *Astrobiology Society* and/or seeking support from existing intergovernmental organizations such as the United Nations Office of Outer Space Affairs.
- Periodic assessment of the totality of evidence for life on specific targets which can be performed by the newly formed *Astrobiology Society* or a group constituted ad-hoc for that task.
- Creating databases (such as the LDKB) of examples of consensus standards in other fields

Verification of results may happen in several stages. Before submitting to a refereed journal, the research group may share data with another research team for a simultaneous publication of findings. After publication, the aforementioned committee may help to build a verification strategy. A curated list of peer-reviewers for life detection claims may be helpful for editors handling journal articles related to the search for life. We recognize that some of these actions require the active participation of journal editors who are part of the scientific community but act under the policies of their journals. Editor and journal engagement is fundamental for the verification protocol. The verification process in any of the described stages would benefit from financial and organizational resources dedicated to this task.

As a group, we assume that there are ethical rules and codes of conduct in our countries and institutions that must be followed during research, publication, reviewing and collaboration that apply to the verification process. Following ethical rules and codes of conduct creates a more trusting environment within the community and favors meaningful collaborations and the verification of results.

This process could facilitate building trust between scientists, the public media, and the general public and increasing public understanding of astrobiology topics. By putting in practice the protocol and verification processes as advisors and teachers we contribute to making good practices a natural way to report, collaborate and verify scientific results (Des Marais, 2008).

5.5 Lessons learned

Communication challenges are not unique to astrobiology. Other fields, such as climate change and pandemic-related studies, have seen a rampant spread of misinformation and disinformation in recent years. Figure 5.2 demonstrates a tool designed by the IPCC to communicate the strength of evidence for a claim as it correlates with the level of agreement in a scientific community that has evaluated the evidence. Public communication of the Assessment Framework should also strive for intuitive, informative explanations. Astrobiology can learn lessons from experts in these other contentious fields to understand how to disseminate clear information and best incentivize a proper balance between claims and caveats in publications.

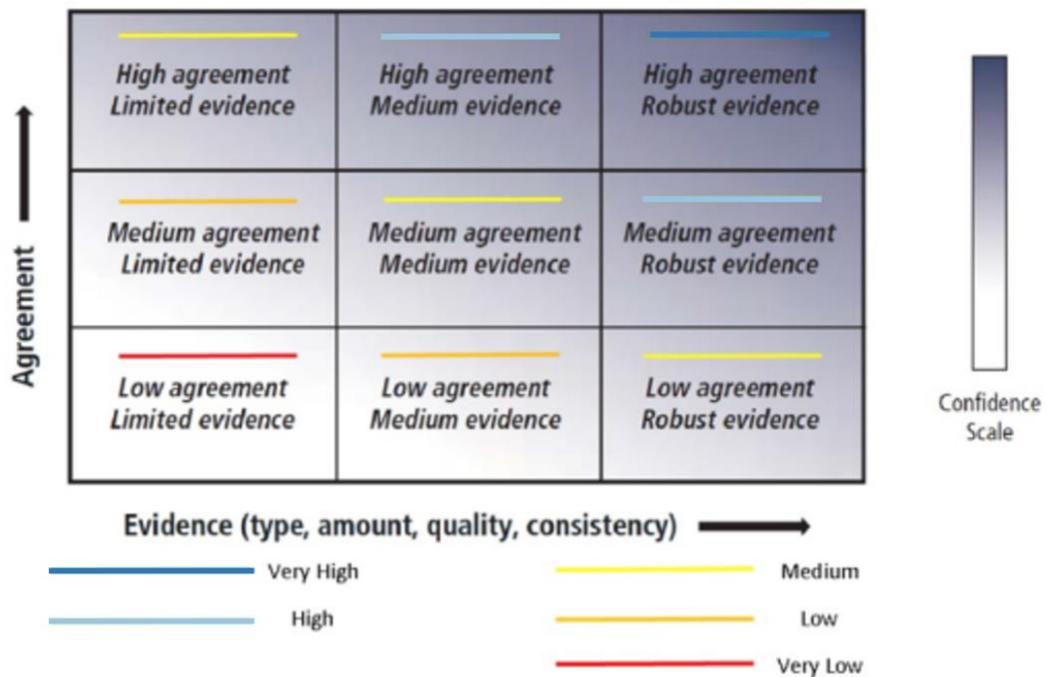


Figure 5.2 Confidence levels are a combination of levels of agreement and evidence. There are five levels shown with colors (after IPCC 2013).

The need to carefully consider wording in announcements is illustrated by the possibly different perceptions of “climate change” and “global warming”, the former being more comprehensive and potentially eliciting more buy-in, but being perceived as less striking (Whitmarsh, 2009; Soutter & Möttus, 2020). Much is

being learned about reporting levels of confidence and the results of scientific investigations still in progress from the reporting on medical consequences of, treatments for, and prevention methods during the ongoing COVID-19 pandemic.

Lessons can be learned from other similar scientific endeavors involving broad community collaborations and/or large taxpayer investments. In planetary defense, a Torino scale (Binzel, 2000) was developed, and refined based on initial reception, to categorize the impact hazard associated with near-Earth small bodies (e.g., asteroids and comets); a topic prone to sensationalization and potentially involving many stakeholders, from scientists, to the broader public, to decision-makers and emergency response agencies. The Torino scale comprises 10 levels that collapse the dimensions of impact energy (consequence) and impact likelihood. In the search for extraterrestrial intelligence (SETI), a Declaration of Principles includes discussion of the transparency of the search, handling of candidate evidence, and reporting procedures (SETI, 2018). Additionally, a Rio scale was developed to convey to the public the significance of a candidate ETI signal (Forgan et al., 2018).

Parallels can also be drawn from the discovery of the Higgs boson at the Large Hadron Collider (LHC). That discovery was years in the making, involved a large international collaboration, and demanded a \approx RY\$6.5b hardware investment (CERN 2021a) similar to that of space missions such as Viking, the James Webb Space Telescope, Mars Sample Return, or life detection mission concepts for ocean worlds (e.g., NASEM 2017). Major LHC results were reported in articles written by collaborations involving hundreds of researchers, rather than the more restricted teams typical of space missions; preprints were reviewed internally (CERN 2021b). The broader scientific community, the media, and the public were kept apprised of progress in the confidence of the detection (e.g., the evolving statistical significance of the detection and uncertainty on the energy of the signal), from the initial shrinking of the search space, to hints of signals, to firmed up information. This was thanks to pre-designed fully open and transparent data sharing with community buy-in, and contrasts with space mission data, which are embargoed for a specific time period, most often until the initial publications have been released. This led to the eventual reports of the Higgs boson discovery being rather anticlimactic, less splashy headlines than an expected conclusion. The major results have not been disputed, and scientific consensus was reached without controversy. For the search for life beyond Earth, identifying and communicating uncertainties is more difficult, as the search space is much less clearly defined, involving many parameters and disciplines. However, this example offers a model for the proposed Assessment Framework “interim report” publications and demonstrates a good example of verification as part of a reporting protocol.

6: Findings and Future High Priority Research Avenues

6.1 Current understanding and findings of the workshop

A consensus for the best approaches to the multiple avenues for future research is critically needed in order to identify and confirm the detection of life elsewhere in the Universe. It is particularly important to accurately portray the search for life as a multi-step process, rather than one achievable with a single investigation or even mission (e.g., Green et al., 2021). Forging a stepwise biosignature Assessment Framework is in the best interest of science and scientists, and can allow us to provide better communication of the significance of new discoveries to the general public. To this end, each new report or planned mission should be evaluated in this context with the realization that even if life is *not* detected, the field will still progress, guided by the systematic use of a biosignature assessment scale. Here we will also address avenues of research identified by the community that should receive particular attention in the context of identifying life elsewhere in the Universe and the steps needed to confirm (or reject) conclusions drawn from collected data.

Enhancing the Flexibility and Power of the Assessment Framework

It has been stressed previously (Green et al., 2021) that there is a continuum from a low to higher certainty level in whether or not a given phenomenon is due to life, and that any additional data can help further our collective understanding. Trust in the veracity of a biosignature claim has two components: (1) uncertainty in the detection due to measurement error and potential contamination and (2) uncertainty in the interpretation due to lack of understanding of abiotic processes and/or an unfamiliar metabolism. Uncertainty in the detection and identification component is more readily quantifiable. Confidence in the interpretation, although harder to quantify, could be increased via improved understanding of environmental context and the likelihood of false negatives and false positives for the potential biosignature under consideration (Green et al., 2021), and further work on the mechanisms of biosignature generation. Nonetheless, although the generalized steps included in the framework do not need to be linear and can be iterative (see Section 2)---depending on measurement constraints and interpretation needs typical to a given field-- it was considered critical to undertake a rigorous study of the signal detection and identification prior to extensive assessment of the signals biogenicity.

Important discussions at the Standards of Evidence workshop addressed the need to identify critical areas that require future research and efforts. Participants across disciplinary divides agreed on the importance of:

- multiple lines of evidence (Assessment Framework Question 3: for a suite of biosignatures, if the detected signal is due to life, we should also see phenomena x and y - do we see x and y?)

- exhaustive examination of contamination and other sources of false positives
- discriminating potential biosignatures from abiotic and prebiotic chemistry
- geological context/habitability
- biosignature preservation/atmospheric lifetime potential

And for topics that both likely require significant additional community discussion to resolve:

- further development of the framework to generate a more detailed set of worked examples that can serve as field-specific guidelines for biosignature assessment
- enhancement of the framework to incorporate statistical methodology at each step, and the development of a ranking or numerical scheme for certainty in biosignature detection and interpretation.

When claiming the detection of an extraterrestrial biosignature, multiple lines of evidence are crucial, in both corroborating a detection, and in ruling out potential contamination. The strongest lines of evidence for life would be multiple measurements of a given biosignature, or even a suite of biosignatures (NASEM, 2019). Previous work has identified multiple different types of biosignatures including: organic molecules, minerals, macro- and micro-structure, chemistry and isotopes (Mustard et al., 2013). Therefore, multiple measurements, and identification of multiple types of potential biosignatures, would provide the strongest evidence of life. These suites of biosignatures can also be strengthened by asking the question “If the detected signal(s) are due to life, we should also see phenomena x and y; do we see x and y?”

Examination of the planetary context and the potential habitability of an environment (see Questions 3 and 4) would help to rule out abiotic processes and strengthen the argument for detection of life. Community work to survey, identify and understand likely abiotic processes that can mimic life’s impact on its environment is a critical preparatory step for applying the Assessment Framework. Participants agreed on the need to undertake laboratory work to better understand abiotic expressions without the imprint of the billions of years of biotic influence we see on Earth, although abiotic processes can be identified and studied via a variety of means. Once identified and characterized, predictions can be made for measurable discriminants that would identify the abiotic process. For example, Question 3 of the Assessment Framework would require researchers to also invert the previous idea and ask “if this signal *isn’t* due to life, but is instead due to the following abiotic process/characteristic, then we should see phenomena p and q”; do we see p and q? This covers concerns vis-a-vis abiotic biomorphs such as the “nanobacteria” observed in meteorite ALH84001 or other systems (McKay et al., 1996; Schieber & Arnett, 2003). Therefore, the tests of the impacts of life on the environment can also be addressed here. One of the ways that these detections can be strengthened is through the use of data from analog environments and experiments (Assessment Framework Questions 3 and 5; e.g., Hays et al., 2017).

Although no analog or experiment is perfect, this work can provide an answer to what we would expect to see if life were present in context. In addition, all else being equal, an analysis of larger samples is preferable - to avoid false negatives due to overlooking sparse life. If we, as a community, assign a high priority to the return of space samples for biosignature investigations on Earth, this will necessitate the development of additional protocols for sample storage, handling, transport, sharing, and curation in the context of international partnerships (Meyer et al., 2020). Application of this framework to the interpretation of potential biosignatures will require the continued development of extensive libraries of known relevant abiotic and biotic signatures for both extant and extinct life (e.g., The Knowledge Base, Chan et al., 2019) and better defining the limits of life to use as standards.

Similarly, the detection of a biosignature in an environment that is known to be habitable for terrestrial-like life would be a line of evidence greatly strengthening that detection, and would be strongly relevant to answering Question 4, on the likelihood that life would produce a signal in this environment, and Question 5 of the Assessment Framework, on whether follow-up experiments could be devised that would help discriminate between a life and non-life hypothesis. However, potential detections of life should not be limited to environments that are known to be habitable for terrestrial life, but should also include potential detections of life as we don't know it.

Another central theme that was raised was the extent that biosignature preservation potential should be considered during target selection, but also as part of the Assessment Framework. Preservation in this sense can refer to a broad range of processes that span disciplines and measurement techniques, from preservation over time in the rock record, to the chemical lifetime and persistence of biosignature gases in a planetary atmosphere.

Given that the workshop included only 3 half days of discussion, several important topics were identified for further discussion in subsequent workshops or other fora. Two of the highest priorities include further work to develop more detailed "worked examples" for specific biosignatures and measurement techniques, and significant further work needed to incorporate statistical methodology into methods to address each Question in the Framework. The workshop participants were able to agree on a fundamental, generalized framework for biosignature assessment, guided by a series of specific questions that could be interpreted for applicability and applied to a broad range of life detection claims. However, specific fields may wish to use the generalized framework to pre-identify and develop the best practices---likely a series of actions and checks---required to address the five framework questions for a particular biosignature and measurement technique. Several of these "worked examples" for different biosignatures and measurement techniques are provided in Section 3, but worked examples not

covered in this section, or more detailed treatments with field-agreed-upon actions may require further community input to develop and adopt.

Another area where discussions were initiated but not finished at the workshop included evolving the Framework to incorporate statistical methodology in answering each Question. There was strong agreement that a systematic understanding of additive uncertainty should be part of any evidence validation activity. Further, workshop participants expressed the need for research that addresses the data and meta-data needs in order to apply statistical methods such as a Bayesian analysis. The value of a Bayesian methodology has previously been endorsed in the Astrobiology Strategy report (NASEM, 2018) since this method incorporates assumptions and prior-information and can then constrain probability of life even without a positive detection (Walker et al., 2018). However, incorporating statistical methodology into the Framework represents a significant amount of future work, but it would help researchers incorporate statistical needs into data collection processes early in a life detection effort. Developing mechanisms to encourage collaboration between biosignature scientists and statistical specialists was identified as a particularly fruitful avenue for enhancing the rigor of biosignature assessment.

Expanding our Definitions of Biosignature Targets

In addition to expanding our knowledge of potential biosignature targets for known metabolisms, potential detections of life as we don't know it would strongly benefit from the contributions of "agnostic biosignatures" (biosignatures that do not rely on life on Earth as an analog or target analyte). All else being equal, an agnostic approach to biosignature search is preferable over a more limited one that includes assumptions about metabolism, ecology, and physiology. An agnostic approach could include theories and applications from information theory and computational simulations (Keefe et al., 2010; Goodwin et al., 2015; Cronin and Walker 2016; Cabrol, 2016). There is less bias from assumptions about context and biological markers, while preserving some of the fundamental laws in nature (i.e., evolution - evolutionary algorithms; entropy - both physical and informational; network dynamics, scalability, etc.). Without adequate understanding, agnostic biosignatures could become a form of false negative, where besides worlds with sparse life and weak signals, we could miss "life in its most general form".

In addition, we should be open to the possibility that chemistry that appears auto-catalytic may not be due to life, but instead may be a form of prebiotic chemistry. An improved understanding of the continuum of chemistry from non-life to life will help us to identify possible false positives for biosignatures. Agnostic biosignatures may allow the detection of planetary environments exhibiting this prebiotic state. Notably, if we posit that evolution of biology from abiotic components is a continuum, we may observe some intermediate stage in which case an agnostic biosignature approach may be even more powerful.

Machine and deep learning (ML/DL) methods are also important to distinguish signal from noise when data is not straightforward, or when there are multiple data

sources. Additionally, a detection implies a probabilistic approach which can include models (i.e., Bayesian) that can be easily adapted to the multiple lines of evidence and include updates on the evidence as well.

Finally, assumptions regarding life detection also need to be explicit. For example, the current working hypothesis for life detections is the *null hypothesis* - that life is not the explanation for the detection. This is an assumption and should be made explicit. Other potential assumptions that could be and often are made include that habitability is for “known life”, rather than life as we don’t know it, or the requirement for surface liquid water for exoplanet habitability. These can be important arguments strengthening the detection of life, but can nonetheless represent assumptions that should be explicitly stated.

6.2 Community Engagement Structures

Given the process and constraints of how contemporary scientific discovery is executed, what might be ways that scientists, journalists, and others involved in science communication can compensate for these structural challenges? (Figure 6.2a)

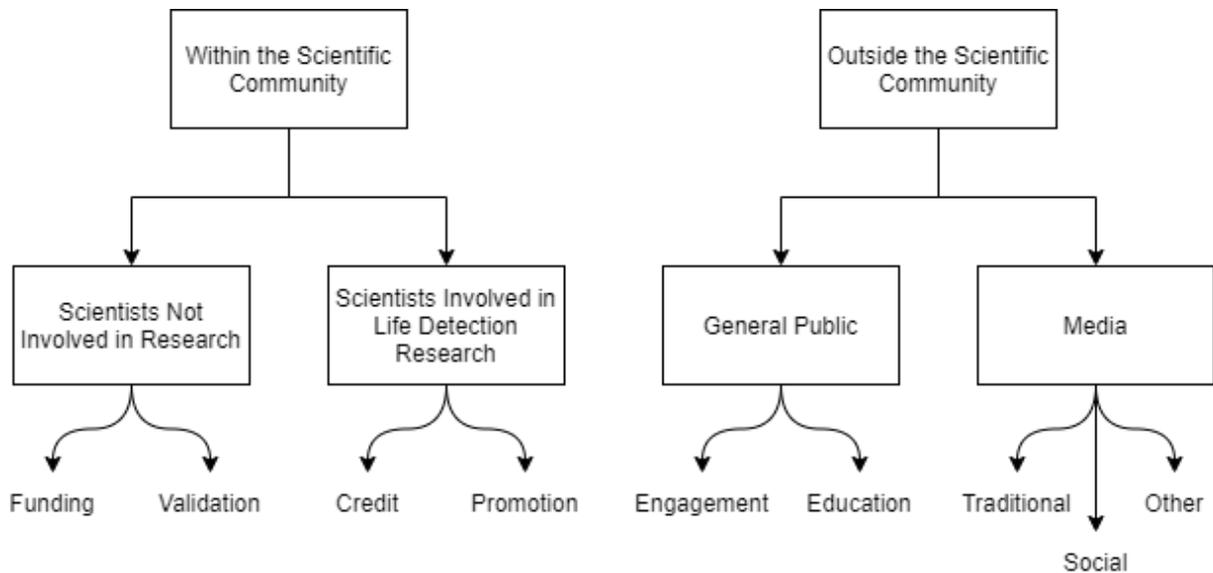


Figure 6.2a. Summary of Stakeholders Involved in Life Detection Findings Dissemination. There are numerous communities, including potentially new ones, that will engage in the scientific discussion of the detection of life elsewhere in the Universe. The squares in this diagram represent stakeholder communities, and the bottom row identifies principal incentives for these communities to be interested in claims of life detection. Scientists are motivated by the possibility of scientific credit, promotion, funding of follow-up work and scientific validation. Outside the scientific community, there is demand for both the engagement and education of the general public as moderated by the media.

As a community, workshop participants came to a few points of consensus on supporting aspects of protocols for biosignature assessment and reporting, some of which may take significant effort to slowly change the status quo. Notably, they fall into roughly four categories including:

- Forming an independent standing international oversight committee to act as a “red team”, rotating members (see below re: current international cooperation relationships).
- Learning from other high-profile research areas such as climate change, LHC/CERN, public health (i.e., vaccinations), epidemiology, etc.
- Encouraging funding support specifically for independent groups to validate data and coordinated publications, including the possibility of pursuing international options for validation.
- Expanding content/access of open servers that host data, pre- and post-publication.

Any and all of these steps could be impactful in combating “disinformation”, as messaging can slip from the control of scientists via the rapid dissemination that comes with a digitally connected population. In particular, incentivizing scientists to collaborate on fewer, more robust, and so higher impact papers could come from the institutions they work at (e.g., universities, federal laboratories, etc.) as well as funding agencies (e.g., NASA, ESA, etc.). If each author/participant involved benefited in some concrete way from the publication(s), then this would in turn incentivize more collaborative, and potentially more rigorous science. However, a mechanism would be needed to distribute the credit for the finding to not only the first author and one particular funding agency, but also the other authors and potentially outside groups validating data.

Funding Mechanisms to Facilitate Collaboration

If we take other constraints into consideration, funding additional, potentially rapid response work to validate (or not) a biosignature, either via alternate analytical methods or by reproducing the results using identical methods used by the original group, would be extremely influential in both confirming the research as well as building trust with the general public. Funding is an obvious place that could serve to transform how our research is done, and the public communication of that research. Examples of a potentially effective mechanism include the short term rapid response funds managed by agencies such as the NSF and the International Ocean Drilling Program that can direct funds towards immediate needs (e.g., in the event of a geohazard). A fund of this kind could be activated towards validation needs in the event of a promising life detection claim.

Several examples also exist of beneficial public-private partnerships for scientific research. In the United States the National Institutes of Health (NIH), and other key governmental agencies such as the FDA and NSF have foundations associated with them (secondary funding support mechanisms) that leverage the agency’s peer review expertise to identify key research. This, coupled with the agencies’ funding management process is used to distribute private funding to targeted individual research as well as to topical research foci. A “Foundation for Astrobiology” with a specific focus on life detection verification could provide an

opportunity for funding from philanthropists to be managed by a government agency and a public board of directors.

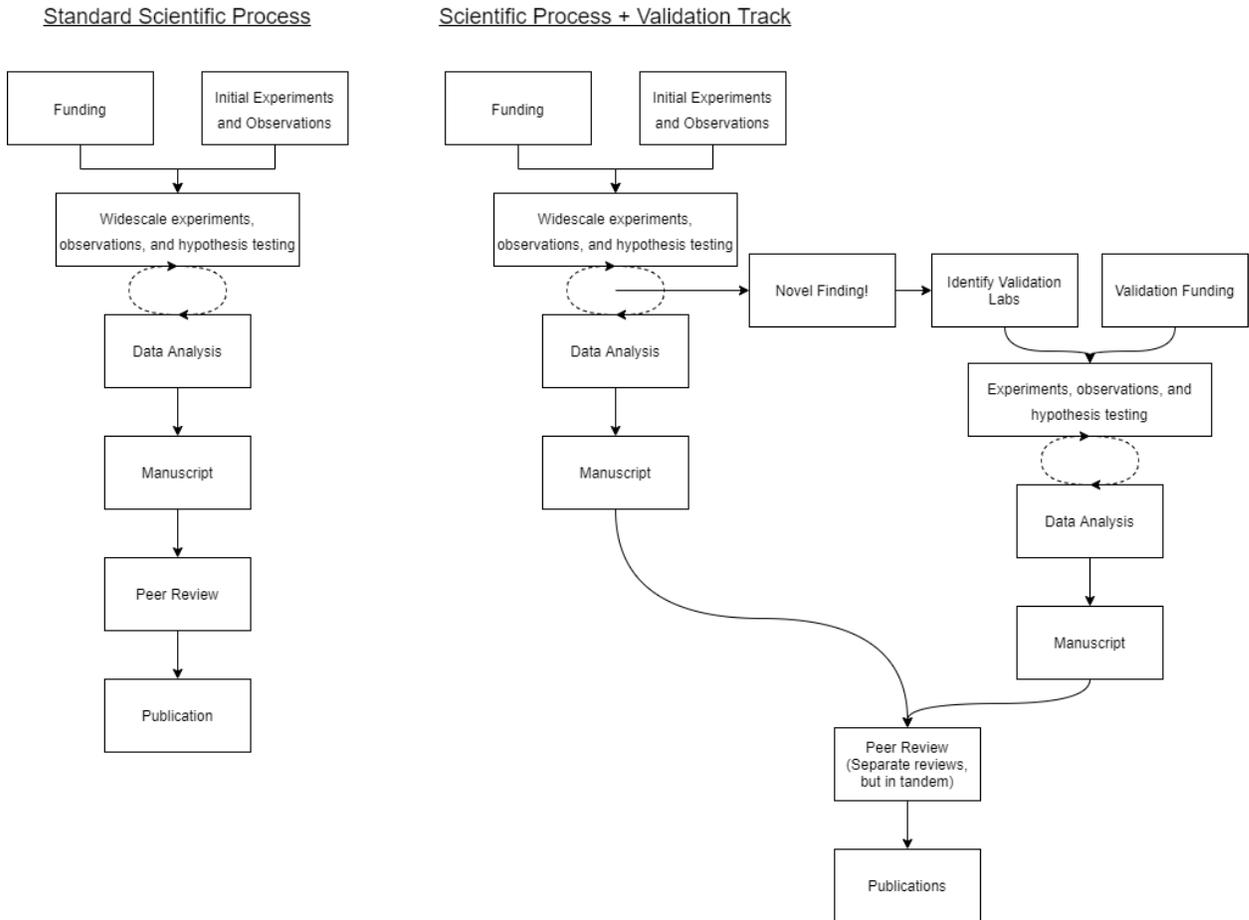


Figure 6.2. *Proposed Workflow for Scientific Discovery Related to Life Detection Incorporating Outside Validation. In the left panel, the standard scientific process that leads to publication follows a predictable pathway. The right panel illustrates a new “validation track” whereby with initial pools of traditional funding, researchers will progress and iterate with increased data, leading to the testing of more precise hypotheses. If life detection is confirmed, funding agencies can then open a wider rapid search for more focused funding for interdisciplinary validation experiments.*

There are other examples of national and international collaborative efforts in a wide range of disciplines, including the establishment of professional societies that might serve as guides for the Astrobiology community, and could be further explored. Existing international collaboration with astrobiology-focused funding responsibilities shared across borders include a number of astrobiology training courses around the world. These include the Iberoamerican School of Astrobiology and the International School of Astrobiology. These are funded by NASA and international governmental funding, including UNESCO. Notably, the former courses are taught in Spanish, thus being deeply inclusive to a significant source of human capital while training a new generation of students in astrobiology. Given the inherently interdisciplinary nature of the search for extraterrestrial life, enlisting

the intellectual and collective assistance of scientific communities around the world may well turn out to be critical for the acceptance of a novel finding. Participants also expressed interest in the formation of an astrobiology professional society that as part of its activities could identify an independent body to contribute to the management of life detection verification and reporting protocol activities, and associated resources.

Incentivizing The Assessment Framework

While workshop participants agreed upon the need to devise a framework for biosignature assessment and reporting, a much more difficult conversation involves how we might incentivize adoption of this framework. We would argue that this framework, if implemented correctly, could be generally applicable to any number of high profile scientific endeavors that will attract a great deal of attention from the scientific community as well as the media and general public. But how can we incentivize the adaptation of existing scientific practices towards a more collective, collaborative response? We wrestle with this question below, and suggest a number of ways that the scientists doing this work, as well as the larger scientific community, could be strongly encouraged to both utilize the framework *and* abide by a principled public discourse. The overarching motivator comes down to funding.

There may be many reasons for scientists not to use this Framework or follow the reporting protocol process without some additional incentives. Many of these incentives, come down to funding and recognition. Funding is what drives science and recognition is critical for continued engagement of funding. Programs to incentivize and reward biosignature validation could follow examples in medicine and social sciences that fund reproducibility studies. While private entities may try to incentivize scientists through high-profile targeted awards focused on risky, “moonshot” research and conspicuous technology demonstrations, participants agreed that transparent, peer-reviewed funding programs have greater promise to provide the necessary background to bolster the Assessment Framework.

Committed funding from the start, for both the research and the validation research (Fig. 6.2), would incentivize following the framework itself, including a reasonable approach to reporting protocols. If as a community, we remove the potential threat of detriment for the original authors (and therefore, a more equally distributed risk to all the authors) this would promote a more rational approach to disseminating the findings. Distributing the reward or recognition of such a groundbreaking discovery would also be more in line with research on climate change (e.g., the IPCC) and work produced by the Large Hadron Collider (i.e., CERN). This diffusion of responsibility and recognition could create a process where each participating scientist (either original research or validation research) shares credit, as well as taking ownership, of their individual contribution, and promote engagement in collaborative research.

References

- Agol, E., Jansen, T., Lacy, B., Robinson, T. D., & Meadows, V. (2015). The Center of Light: Spectroastrometric Detection of Exomoons. *The Astrophysical Journal*, 812(1), 5. <https://doi.org/10.1088/0004-637X/812/1/5>
- Airapetian, V. S., Gloer, A., Khazanov, G. v., Loyd, R. O. P., France, K., Sojka, J., Danchi, W. C., & Liemohn, M. W. (2017). How Hospitable Are Space Weather Affected Habitable Zones? The Role of Ion Escape. *The Astrophysical Journal*, 836(1), L3. <https://doi.org/10.3847/2041-8213/836/1/L3>
- Airapetian, V. S., Jackman, C. H., Mlynczak, M., Danchi, W., & Hunt, L. (2017). Atmospheric Beacons of Life from Exoplanets Around G and K Stars. *Scientific Reports*, October, 1–9. <https://doi.org/10.1038/s41598-017-14192-4>
- Akins, A.B., Lincowski, A.P., Meadows, V.S. and Steffes, P.G. (2021). Complications in the ALMA Detection of Phosphine at Venus. *The Astrophysical Journal Letters*, 907(2), p.L27.
- Albers, Casper J., Henk AL Kiers, Don van Ravenzwaaij, and Victoria Savalei. "Credible confidence: A pragmatic view on the frequentist vs Bayesian debate." *Collabra: Psychology* 4, no. 1 (2018).
- Alleon, J., Flannery, D.T., Ferrali, N., Williford, K.H., Zhang, Y., Schuessler, J.A. & Summons, R.E. (2019). Organo-mineral associations in chert of the 3.5 Ga Mount Ada Basalt raise questions about the origin of organic matter in Paleoarchean hydrothermally influenced sediment. *Nature Scientific Reports*, 9, 16712. <https://doi.org/10.1038/s41598-019-53272-5>
- Allwood, A. C., Grotzinger, J. P., Knoll, A. H., Burch, I. W., Anderson, M. S., Coleman, M. L., & Kanik, I. (2009). Controls on development and diversity of Early Archean stromatolites. *Proceedings of the National Academy of Sciences*, 106(24), 9548–9555. <https://doi.org/10.1073/pnas.0903323106>
- Allwood, A. C., Kamber, B. S., Walter, M. R., Burch, I. W. & Kanik, I. (2010). Trace elements record depositional history of an Early Archean stromatolitic carbonate platform. *Chem. Geol.* 270, 148–163.
- Allwood, A. C., Rosing, M. T., Flannery, D. T., Hurowitz, J. A., & Heirweh, C. M. (2018). Reassessing evidence of life in 3,700-million-year-old rocks of Greenland. *Nature*, 563(7730), 241–244. <https://doi.org/10.1038/s41586-018-0610-4>
- Allwood, A. C., Walter, M. R., Kamber, B. S., Marshall, C. P., & Burch, I. W. (2006). Stromatolite reef from the Early Archaean era of Australia. *Nature*, 441, 714–718. <https://doi.org/10.1038/nature04764>
- Awramik, S. M., & Grey, K. (2005). Stromatolites: Biogenicity, biosignatures, and bioconfusion (R. B. Hoover, G. V. Levin, A. Y. Rozanov, & G. R. Gladstone, Eds.; p. 59060P). <https://doi.org/10.1117/12.625556>
- Barlow, E., Van Kranendonk, M. J., Yamaguchi, K. E., Ikehara, M., & Lepland, A. (2016). Lithostratigraphic analysis of a new stromatolite–thrombolite reef from across the rise of atmospheric oxygen in the Paleoproterozoic Turee Creek Group, Western Australia. *Geobiology*, 14(4), 317–343. <https://doi.org/10.1111/gbi.12175>
- Batchelor, M.T., Burne, R.V., Henry, B.I. & Watt, S.D. (2003) Mathematical and image analysis of stromatolite morphogenesis. *Mathematical Geology*, 35, 789-803.
- Beichman, Charles, Bjoern Benneke, Heather Knutson, Roger Smith, Pierre-Olivier Lagage, Courtney Dressing, David Latham et al. "Observations of transiting exoplanets with the James Webb Space Telescope (JWST)." *Publications of the Astronomical Society of the Pacific* 126, no. 946 (2014): 1134.
- Binzel, R.P. (2000) The Torino impact hazard scale. *Planetary and Space Science* 48(4), 297-303.

- Bishop, Dorothy VM (1990). "How to increase your chances of obtaining a significant association between handedness and disorder." *Journal of clinical and experimental neuropsychology* 12, no. 5: 812-816.
- Bishop, Dorothy (2019). "Rein in the four horsemen of irreproducibility." *Nature* 568, no. 7753: 435-436.
- Bosak, T., Knoll, A. H., & Petroff, A. P. (2013). The Meaning of Stromatolites. *Annual Review of Earth and Planetary Sciences*, 41(1), 21–44. <https://doi.org/10.1146/annurev-earth-042711-105327>
- Brandt, T. D., & Spiegel, D. S. (2014). Prospects for detecting oxygen, water, and chlorophyll on an exo-Earth [Earth and Planetary Astrophysics; Instrumentation and Methods for Astrophysics]. *Proceedings of the National Academy of Sciences*, 111(37), 13278–13283. <https://doi.org/10.1073/pnas.1407296111>
- Brasier, Martin D., Owen R. Green, Andrew P. Jephcoat, Annette K. Kleppe, Martin J. Van Kranendonk, John F. Lindsay, Andrew Steele, and Nathalie V. Grassineau (2002). "Questioning the evidence for Earth's oldest fossils." *Nature* 416, no. 6876: 76-81.
- Brocks, Jochen J., Graham A. Logan, Roger Buick, and Roger E. Summons (1999). "Archean molecular fossils and the early rise of eukaryotes." *Science* 285, no. 5430: 1033-1036.
- Buick, R., Dunlop, J. S. R., & Groves, D. I. (1981). Stromatolite recognition in ancient rocks: An appraisal of irregularly laminated structures in an Early Archaean chert-barite unit from North Pole, Western Australia. *Alcheringa: An Australasian Journal of Palaeontology*, 5(3), 161–181. <https://doi.org/10.1080/03115518108566999>
- Burton, A.S, Stern, J.C., Elsil, J.E., Glavin D.P. & Dworkin, J.P. (2012) Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. *Chemical Society Reviews*, 41, 5459-5472.
- Byerly, G. R., Lowe, D. R. & Walsh, M. M. (1986). Stromatolites from the 3,300–3,500 Myr Swaziland Supergroup, Barberton Mountain Land, South Africa. *Nature* 319, 489–491.
- Cable, Morgan L., et al. (2016). "Enceladus Life Finder: The search for life in a habitable moon.".
- Catling, D. C., & Claire, M. W. (2005). How earth's atmosphere evolved to an oxic state: A status report. In *Earth and Planetary Science Letters* (Vol. 237, Issues 1–2, pp. 1–20). <https://doi.org/10.1016/j.epsl.2005.06.013>
- Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., Crisp, D., Robinson, T. D., DasSarma, S., Rushby, A. J., del Genio, A., Bains, W., & Domagal-Goldman, S. (2018). Exoplanet Biosignatures: A Framework for Their Assessment. *Astrobiology*, 18(6), 709–738. <https://doi.org/10.1089/ast.2017.1737>
- CERN (2021a) Facts and figures about the LHC. <https://home.cern/resources/faqs/facts-and-figures-about-lhc>. Retrieved July 2021.
- CERN (2021b) Experimental Physics Publishing Policy. <https://ep-dep.web.cern.ch/content/ep-publishing-policy>. Retrieved Sept. 2021.
- Chafetz, H. S., & Guidry, S. A. (1999). Bacterial shrubs, crystal shrubs, and ray-crystal shrubs: Bacterial vs. abiotic precipitation. *Sedimentary Geology*, 126(1–4), 57–74. [https://doi.org/10.1016/S0037-0738\(99\)00032-9](https://doi.org/10.1016/S0037-0738(99)00032-9)
- Chan, M. A., Hinman, N. W., Potter-McIntyre, S. L., Schubert, K. E., Gillams, R. J., Awramik, S. M., Boston, P.J., Bower, D.M, Des Marais, D.J., Farmer, D.J., Jia, T.Z., King, P.L., Hazen, R.M., Léveillé, R.J., Papineau, D., Rempfert, K.R., Sánchez-Román, M., Spear, J.R., Southam, G., Stern, J.C., Cleaves, H. J. (2019). Deciphering biosignatures in planetary contexts. *Astrobiology*, 19(9), 1075-1102.
- Chassefière, E. (1996). Hydrodynamic escape of oxygen from primitive atmospheres: applications to the cases of Venus and Mars. *Icarus*, 552, 537–552. <https://doi.org/10.1006/icar.1996.0229>
- Chassefière, E., Wieler, R., Marty, B., & Leblanc, F. (2012). The evolution of Venus: Present state of knowledge and future exploration. *Planetary and Space Science*, 63–64, 15–23. <https://doi.org/10.1016/j.pss.2011.04.007>

- Cowan, N. B., Agol, E., Meadows, V. S., Robinson, T., Livengood, T. a., Deming, D., Lisse, C. M., A'Hearn, M. F., Wellnitz, D. D., Seager, S., & Charbonneau, D. (2009). Alien Maps of an Ocean-Bearing World. *The Astrophysical Journal*, 700(2), 915–923. <https://doi.org/10.1088/0004-637X/700/2/915>
- Cronin L. & Walker S.I. (2016). Beyond prebiotic chemistry. *Science*, 352(6290):1174–1175.
- Cuerno, R., Escudero, C., García-Ruiz, J.M. & Herrero, M.A. (2011) Pattern formation in stromatolites: insights from mathematical modelling. *Journal of the Royal Society Interface*, 9(70), 1051-1062.
- Defrère, D., Léger, A., Absil, O., Beichman, C., Biller, B., Danchi, W. C., Ergenzinger, K., Eiroa, C., Ertel, S., Fridlund, M., Muñoz, A. G., Gillon, M., Glasse, A., Godolt, M., Grenfell, J. L., Kraus, S., Labadie, L., Lacour, S., Liseau, R., ... Wyatt, M. (2018). Space-based infrared interferometry to study exoplanetary atmospheres. *Experimental Astronomy*, 46(3), 543–560. <https://doi.org/10.1007/s10686-018-9613-2>
- Des Marais, D. J., Harwit, M. O., Jucks, K. W., Kasting, J. F., Lin, D. N. C. C., Lunine, J. I., Schneider, J., Seager, S., Traub, W. A., & Woolf, N. J. (2002). Remote Sensing of Planetary Properties and Biosignatures on Extrasolar Terrestrial Planets. *Astrobiology*, 2(2), 153–181. <https://doi.org/10.1089/15311070260192246>.
- Des marais, D.J., Nuth, III, J.A., Allamandola, L.J., Boss, A.P., Farmer, J.D., Hoehler, T.M., Jakosky, B.M., Meadows, V.S., Pohorille, A., Runnegar, B. & Spormann, A.M. (2008). The NASA astrobiology roadmap. *Astrobiology*, 8(4), 715-730.
- Djokic, T., Van Kranendonk, M., Campbell, K. et al. Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nat Commun* 8, 15263 (2017). <https://doi.org/10.1038/ncomms15263>
- Domagal-Goldman, S. D., Segura, A., Claire, M. W., Robinson, T. D., & Meadows, V. S. (2014). Abiotic Ozone and Oxygen in Atmospheres Similar to Prebiotic Earth [Earth and Planetary Astrophysics]. *The Astrophysical Journal*, 792(90), 43. <https://doi.org/10.1088/0004-637X/792/2/90>
- Farrance, Ian, and Robert Frenkel (2012). "Uncertainty of measurement: a review of the rules for calculating uncertainty components through functional relationships." *The Clinical Biochemist Reviews* 33, no. 2: 49.
- Feng, Y. K., Robinson, T. D., Fortney, J. J., Lupu, R. E., Marley, M. S., Lewis, N. K., Macintosh, B., & Line, M. R. (2018). Characterizing Earth Analogs in Reflected Light: Atmospheric Retrieval Studies for Future Space Telescopes. *The Astronomical Journal*, 155(5), 200. <https://doi.org/10.3847/1538-3881/aab95c>
- Flannery, D. T., Allwood, A. C., & Van Kranendonk, M. J. (2016). Lacustrine facies dependence of highly ¹³C-depleted organic matter during the global age of methanotrophy. *Precambrian Research*, 285, 216–241. <https://doi.org/10.1016/j.precamres.2016.09.021>
- Flannery, D. T., & Walter, M. R. (2011). Archean tufted microbial mats and the Great Oxidation Event: New insights into an ancient problem. *Australian Journal of Earth Sciences*, 59(1), 1–11. <https://doi.org/10.1080/08120099.2011.607849>
- Flügel, E., & Munnecke, A. (2010). Microfacies of carbonate rocks: Analysis, interpretation and application (2nd ed). *Springer*.
- Forgan et al. (2018) Rio 2.0: revising the Rio scale for SETI detections. *Int. J. Astrobio.* 18, 336-344. <https://doi.org/10.1017/S1473550418000162>
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... & Barabási, A. L. (2018). Science of science. *Science*, 359(6379).
- French, Katherine L., Christian Hallmann, Janet M. Hope, Petra L. Schoon, J. Alex Zumberge, Yosuke Hoshino, Carl A. Peters et al. (2015). "Reappraisal of hydrocarbon biomarkers in Archean rocks." *Proceedings of the National Academy of Sciences* 112, no. 19: 5915-5920.

- Frick, A., Mogul, R., Stabekis, P., Conley, C. A., & Ehrenfreund, P. (2014). Overview of current capabilities and research and technology developments for planetary protection. *Advances in Space Research*, 54(2), 221-240.
- Gao, P., Hu, R., Robinson, T. D., Li, C., & Yung, Y. L. (2015). Stability of CO₂ Atmospheres on Dessicated M Dwarf Exoplanets. *The Astrophysical Journal*, 806(2), 249. <https://doi.org/10.1088/0004-637X/806/2/249>
- Gates, D. M., Keegan, H. J., Schleter, J. C., & Weidner, V. R. (1965). Spectral Properties of Plants. *Applied Optics*, 4(1), 11. <https://doi.org/10.1364/AO.4.000011>
- Gaudi, B. S., Seager, S., Mennesson, B., Kiessling, A., Warfield, K., Cahoy, K., Clarke, J. T., Domagal-Goldman, S., Feinberg, L., Guyon, O., Kasdin, J., Mawet, D., Plavchan, P., Robinson, T., Rogers, L., Scowen, P., Somerville, R., Stapelfeldt, K., Stark, C., ... Zellem, R. (2020). The Habitable Exoplanet Observatory (HabEx) Mission Concept Study Final Report. *ArXiv Preprint 2001.06683*. <http://arxiv.org/abs/2001.06683>
- Gillmann, C., Chassefière, E., & Lognonné, P. (2009). A consistent picture of early hydrodynamic escape of Venus atmosphere explaining present Ne and Ar isotopic ratios and low oxygen atmospheric content. *Earth and Planetary Science Letters*, 286(3–4), 503–513. <https://doi.org/10.1016/j.epsl.2009.07.016>.
- Glavin, D.P. & Dworkin, J.P. (2009) Enrichment of the amino acid L-isovaline by aqueous alteration on CI and CM meteorite parent bodies. *Proceedings of the National Academy of Sciences*, 106, 5487-5492.
- Glavin, D.P., Callahan, M.P., Dworkin, J.P., & Elsila, J.E. (2010) The effects of parent body processes on amino acids in carbonaceous chondrites. *Meteoritics and Planetary Science*, 45, 1948-1972.
- Glavin, D.P., Elsila, J.E., Burton, A.S., Callahan, M.P., Dworkin, J.P., Hilts, R.W. and Herd, C.D.K. (2012). Unusual non-terrestrial L-proteinogenic amino acid excesses in the Tagish Lake meteorite. *Meteoritics and Planetary Science*, 47, 1347-1364.
- Glavin, D.P., Freissinet, C., Miller, K.E., Eigenbrode, J.L., Brunner, A.E., Buch, A., Sutter B. Archer, P.D. Jr., Atreya, S.K., Brinckerhoff, W.B., Cabane, M., Coll, P., Conrad, P.G., Coscia, D., Dworking, J.P., Franz, H.B., Grotzinger, J.P., Leshin, L.A., Martin, M.G., Mckay, C., Ming, D.W., Navarro-González, R., Pavlov, A., Steele, A., Summons, R.E., Szopa, C., Teinturier, S. & Mahaffy, P.R. (2013). Evidence for perchlorates and the origin of chlorinated hydrocarbons detected by SAM at the Rocknest aeolian deposit in Gale Crater. *Journal of Geophysical Research: Planets* 118(10), 1955-1973.
- Glavin, D.P., Burton, A.S., Elsila, J.E., Aponte, J.C. & Dworkin, J.P. (2020) The search for chiral asymmetry as a potential biosignature in our solar system. *Chemical Reviews*, 120, 4660-4689.
- Goodwin S., Gade A.M., Byrom M., Herrera B., Spears C., Anslyn E.V., & Ellington A.D. (2015). Next-generation sequencing as input for chemometrics in differential sensing routines. *Angewandte Chemie*, 127:6437–6440.
- Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J.-M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M.-A. H., ... Zak, E. J. (2017). The HITRAN 2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 130, 4–50. <https://doi.org/10.1016/j.jqsrt.2017.06.038>
- Greaves, J.S., Richards, A., Bains, W., Rimmer, P.B., Sagawa, H., Clements, D.L., Seager, S., Petkowski, J.J., Sousa-Silva, C., Ranjan, S. and Drabek-Maunder, E. (2021). Phosphine gas in the cloud decks of Venus. *Nature Astronomy*, 5(7), pp.655-664.
- Green, J., Hoehler, T., Neveu, M., Domagal-Goldman, S., Scalice, D., and Voytek, M. Call for a framework for reporting evidence for life beyond Earth. *Nature* 598, 575–579 (2021). <https://doi.org/10.1038/s41586-021-03804-9>

- Grey, K., & Awramik, S. (2020). Handbook for the study and description of microbialites. *Geological Survey of Western Australia Bulletin*, vol. 147
<https://dmpbookshop.eruditetechnologies.com.au/product/microbialites.do>
- Grotzinger, J. P., & Rothman, D. H. (1996). An abiotic model for stromatolite morphogenesis. *Nature*, 383(6599), 423–425. <https://doi.org/10.1038/383423a0>
- Hand, K.P., Murray, A.E., Garvin, J., Horst, S., Brinkerhoff, W., Edgett, K., Hoehler, T., Russell, M., Rhoden, A., Yingst, R.A., German, C., Schmidt, B., Paranicas, C., Smith, D., Willis, P., Hayes, A., Ehlmann, B., Lunine, J., Templeton, A., Neelson, K., Christner, B., Cable, M., Craft, K., Pappalardo, R., and Phillips, C. (2017) Science goals, objectives, and investigations of the 2016 Europa Lander Science Definition Team Report [abstract 2492]. In *48th Lunar and Planetary Science Conference*, Lunar and Planetary Institute, Houston, LPI Contribution No. 1964.
- Harman, C. E., Schwieterman, E. W., Schottelkotte, J. C., & Kasting, J. F. (2015). Abiotic O₂ Levels on Planets Around F, G, K, and M Stars: Possible False Positives for Life? *The Astrophysical Journal*, 812(2), 137. <https://doi.org/10.1088/0004-637X/812/2/137>.
- Hayes, J.M., Freeman, K.H., Popp, B.N. & Hoham, C.H. (1990) Compounds-specific isotopic analyses: A novel tool for reconstruction of ancient biogeochemical processes. *Organic Geochemistry*, 16, 1115-1128.
- Hays, L.E., Graham, H.V., Des Marais, D.J., Hausrath, E.M., Horgan, B., McCollom, T.M., Parenteau, M.N., Potter-McIntyre, S.L., Williams, A.J., Lynch, K.L., (2017). Biosignature Preservation and Detection in Mars Analog Environments. *Astrobiology* 17, 363-400.
- Hecht, M.H., Kounaves, S.P., Quinn, R.C., West, S.J., Young, S.M.M., Ming, D.W., Catling, D.C., Clark, B.C., Boynton, W.V., Hoffman, J. and DeFlores, L.P., (2009). Detection of perchlorate and the soluble chemistry of martian soil at the Phoenix lander site. *Science*, 325(5936), pp.64-67.
- Hedges, J.I. & Keil, R.G. (1995) Sedimentary organic matter preservation: an assessment and speculative synthesis. *Marine Chemistry*, 49, 81-115.
- Hendrix, Amanda R., et al. (2019). "The NASA roadmap to ocean worlds." *Astrobiology* 19.1, 1-27.
- Hegde, S., Paulino-Lima, I. G., Kent, R., Kaltenecker, L., & Rothschild, L. (2015). Surface biosignatures of exo-Earths: Remote detection of extraterrestrial life. *Proceedings of the National Academy of Sciences*, 112(13), 3886–3891.
<https://doi.org/10.1073/pnas.1421237112>
- Hickman, A. H., Van Kranendonk, M. J., & Grey, K. (2011). State Geoheritage Reserve R50149 (Trendall Reserve), North Pole, Pilbara Craton, Western Australia: Geology and Evidence for Early Archean Life (pp. 1-32). *Geological Survey of Western Australia*.
- Hickman-Lewis, K., Cavalazzi, B., Foucher, F. & Westall F. (2018) Most ancient evidence for life in the Barberton greenstone belt: Microbial mats and biofabrics of the 3.47 Ga Middle Marker horizon. *Precambrian Research*, 312, 45-67.
- Hickman-Lewis, K., Gautret, P., Arbaret, L., Sorieul, S., De Wit, R., Foucher, F., Cavalazzi, B. & Westall, F. (2019) Mechanistic morphogenesis of organo-sedimentary structures growing under geochemically stressed conditions: Keystone to proving the biogenicity of some Archean stromatolites? *Geosciences*, 9(8), 359. <https://doi.org/10.3390/geosciences9080359>.
- Hidalgo, S. L., Pietrinferni, A., Cassisi, S., Salaris, M., Mucciarelli, A., Savino, A., Aparicio, A., Aguirre, V. S., & Verma, K. (2018). The Updated BaSTI Stellar Evolution Models and Isochrones. I. Solar-scaled Calculations. *The Astrophysical Journal*, 856(2), 125.
<https://doi.org/10.3847/1538-4357/aab158>
- Hoehler TM. (2007). An energy balance concept for habitability. *Astrobiology*, 7(6):824-38. doi: 10.1089/ast.2006.0095. PMID: 18163865.

- Hong, Y.-C., Raymond, S. N., Nicholson, P. D., & Lunine, J. I. (2018). Innocent Bystanders: Orbital Dynamics of Exomoons During Planet–Planet Scattering. *The Astrophysical Journal*, 852(2), 85. <https://doi.org/10.3847/1538-4357/aaa0db>.
- Homann, M., Sansjofre, P., Van Zuilen, M., Heubeck, C., Gong, J., Killingsworth, B., Foster, I.S., Airo, A., Van Kranendonk, M.J., Ader, M. & Lalonde, S.V. (2018) Microbial life and biogeochemical cycling on land 3,220 million years ago. *Nature Geoscience*, 11, 665-671.
- Hu, R., Peterson, L., & Wolf, E. T. (2020). O₂- and CO-rich Atmospheres for Potentially Habitable Environments on TRAPPIST-1 Planets. *The Astrophysical Journal*, 888(2), 122. <https://doi.org/10.3847/1538-4357/ab5f07>
- Hu, R., Seager, S., & Bains, W. (2012). Photochemistry in Terrestrial Exoplanet Atmospheres. I. Photochemistry Model and Benchmark Cases. *The Astrophysical Journal*, 761(2), 166. <https://doi.org/10.1088/0004-637X/761/2/166>
- Javaux, E.J. (2019). Challenges in evidencing the earliest traces of life. *Nature* 572, 451–460. <https://doi.org/10.1038/s41586-019-1436-4>
- Johnson SS. (2018a) “Agnostic Approaches to Life Detection.”. 2018. *Presentation to the Committee on An Astrobiology Science Strategy for Search for Life in the Universe*, Irvine, CA, January 18.
- Johnson SS, Anslyn EV, Graham HV, Mahaffy PR, Ellington AD. Fingerprinting Non-Terran Biosignatures. *Astrobiology*, 18(7):915–922.
- Johnson SS, Graham H, Anslyn E, Conrad P, Cronin L, Ellington A, Elsila J, et al. (2018). “Agnostic Biosignatures: Towards a More Inclusive Life Detection Strategy.”. *White paper submitted to the Committee on an Astrobiology Science Strategy for the Search for Life in the Universe*.
- Kaltenegger, L. (2017). How to Characterize Habitable Worlds and Signs of Life. *Annual Review of Astronomy and Astrophysics*, 55(1), 433–485. <https://doi.org/10.1146/annurev-astro-082214-122238>.
- Karman, T., Gordon, I. E., van der Avoird, A., Baranov, Y. I., Boulet, C., Drouin, B. J., Groenenboom, G. C., Gustafsson, M., Hartmann, J.-M., Kurucz, R. L., Rothman, L. S., Sun, K., Sung, K., Thalman, R., Tran, H., Wishnow, E. H., Wordsworth, R., Vigasin, A. A., Volkamer, R., & van der Zande, W. J. (2019). Update of the HITRAN collision-induced absorption section. *Icarus*, 328(October 2018), 160–175. <https://doi.org/10.1016/j.icarus.2019.02.034>
- Kasting, J. F., Kopparapu, R., Ramirez, R. M., & Harman, C. E. (2014). Remote life-detection criteria, habitable zone boundaries, and the frequency of Earth-like planets around M and late K stars. *Proceedings of the National Academy of Sciences*, 111(35), 12641–12646. <https://doi.org/10.1073/pnas.1309107110>
- Kasting, J. F., & Pollack, J. B. (1983). Loss of water from Venus. I. Hydrodynamic escape of hydrogen. *Icarus*, 53(3), 479–508. [https://doi.org/10.1016/0019-1035\(83\)90212-9](https://doi.org/10.1016/0019-1035(83)90212-9)
- Kawashima, Y., & Rugheimer, S. (2019). Theoretical Reflectance Spectra of Earth-like Planets through Their Evolutions: Impact of Clouds on the Detectability of Oxygen, Water, and Methane with Future Direct Imaging Missions. *The Astronomical Journal*, 157(5), 213. <https://doi.org/10.3847/1538-3881/ab14e3>
- Keefe AD, Pai S, Ellington A. (2010). Aptamers as therapeutics. *Nature Reviews Drug Discovery*, 9:537–550.
- Kiang, N. Y., Siefert, J., Govindjee, & Blankenship, R. E. (2007). Spectral signatures of photosynthesis I: Review of Earth organisms. *Astrobiology, Special Issue on M Stars*, 7(1), 222–251.
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., Eymet, V., Robinson, T. D., Mahadevan, S., Terrien, R. C., Domagal-Goldman, S., Meadows, V., & Deshpande, R. (2013). Habitable zones around main-sequence stars: New estimates, *The Astrophysical Journal*, 765(2), 131. <https://doi.org/10.1088/0004-637X/765/2/131>

- Krissansen-Totton, J., Bergsman, D. S., & Catling, D. C. (2016). On Detecting Biospheres from Chemical Thermodynamic Disequilibrium in Planetary Atmospheres. *Astrobiology*, 16(1), 39–67. <https://doi.org/10.1089/ast.2015.1327>
- Krissansen-Totton, J., Fortney, J. J., Nimmo, F., & Wogan, N. (2021). Oxygen False Positives on Habitable Zone Planets Around Sun-Like Stars. *AGU Advances*, 2(2). <https://doi.org/10.1029/2020AV000294>
- Krissansen-Totton, J., Olson, S., & Catling, D. C. (2018). Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life. *Science Advances*, 4(1), eaao5747. <https://doi.org/10.1126/sciadv.aao5747>
- Lammer, H., Sproß, L., Grenfell, J. L., Scherf, M., Fossati, L., Lendl, M., & Cubillos, P. E. (2019). The Role of N₂ as a Geo-Biosignature for the Detection and Characterization of Earth-like Habitats. *Astrobiology*, 19(7), 927–950. <https://doi.org/10.1089/ast.2018.1914>
- Lanier, W. P. (1986). Approximate growth rates of Early Proterozoic microstromatolites as deduced by biomass productivity. *Palaios*, 525-542. <https://doi.org/10.2307/3514705>
- Lehmer, O. R., Catling, D. C., Parenteau, M. N., & Hoehler, T. M. (2018). The Productivity of Oxygenic Photosynthesis around Cool, M Dwarf Stars. *The Astrophysical Journal*, 859(2), 171. <https://doi.org/10.3847/1538-4357/aac104>
- Lehmer, O. R., Catling, D. C., Parenteau, M. N., Kiang, N. Y., & Hoehler, T. M. (2021). The peak absorbance wavelength of photosynthetic pigments around other stars from spectral optimization. *Frontier in Astronomy and Space Sciences*, 8(July), 1–10. <https://doi.org/10.3389/fspas.2021.689441>
- Lepot, K. (2020). Signatures of early microbial life from the Archean (4 to 2.5 Ga) eon. *Earth-Science Reviews*, 209, 103296. <https://doi.org/10.1016/j.earscirev.2020.103296>
- Levin, G.V. & P.A. Straat (1976). Viking labeled release biology experiment: interim results. *Science*, 194(4271), 1322-1329.
- Levin, G.V., & P.A. Straat (2016). The case for extant life on Mars and its possible detection by the Viking labeled release experiment. *Astrobiology* 16, no. 10, 798-810.
- Line, M. R., Quanz, S. P., Schwieterman, E. W., Fortney, J. J., Stevenson, K. B., Greene, T., Zellem, R., Morley, C., Kataria, T., Tremblay, L., Menneson, B., Iyer, A., Mawet, D., Iro, N., Kaltenegger, L., & Defrere, D. (2019). The Importance of Thermal Emission Spectroscopy for Understanding Terrestrial Exoplanets. *Bulletin of the American Astronomical Society*, 51(3), 271. <https://ui.adsabs.harvard.edu/abs/2019BAAS...51c.271L/abstract>
- López-Morales, M., Currie, T., Teske, J., Gaidos, E., Kempton, E., Males, J., Lewis, N., Rackham, B.V., Ben-Ami, S., Birkby, J. and Charbonneau, D. (2019). Detecting earth-like biosignatures on rocky exoplanets around nearby stars with ground-based extremely large telescopes. *arXiv preprint arXiv:1903.09523*.
- Lovelock, J. E., & Kaplan, I. R. (1975). Thermodynamics and the Recognition of Alien Biospheres [and Discussion]. *Proceedings of the Royal Society B: Biological Sciences*, 189(1095), 167–181. <https://doi.org/10.1098/rspb.1975.0051>
- Lowe, D. R. (1980). Stromatolites 3,400-Myr old from the Archean of Western Australia. *Nature*, 284(5755), 441–443. <https://doi.org/10.1038/284441a0>
- Lowe, D. R. (1994). Abiological origin of described stromatolites older than 3.2 Ga. *Geology*, 22, 387–390.
- Lloyd, R. O. P., France, K., Youngblood, A., Schneider, C., Brown, A., Hu, R., Linsky, J., Froning, C. S., Redfield, S., Rugheimer, S., & Tian, F. (2016). The Muscles Treasury Survey. III. X-Ray to Infrared Spectra of 11 M and K Stars Hosting Planets. *The Astrophysical Journal*, 824(2), 102. <https://doi.org/10.3847/0004-637X/824/2/102>
- Luger, R., & Barnes, R. (2015). Extreme Water Loss and Abiotic O₂ Buildup on Planets Throughout the Habitable Zones of M Dwarfs. *Astrobiology*, 15(2), 119–143. <https://doi.org/10.1089/ast.2014.1231>

- LUVOIR Team, (2019). The LUVOIR mission concept study final report. *arXiv preprint arXiv:1912.06219*.
- Lustig-Yaeger, J., Meadows, V. S., Tovar Mendoza, G., Schwieterman, E. W., Fujii, Y., Luger, R., & Robinson, T. D. (2018). Detecting Ocean Glint on Exoplanets Using Multiphase Mapping. *The Astronomical Journal*, 156(6), 301. <https://doi.org/10.3847/1538-3881/aaed3a>
- MacKenzie, Shannon M., et al. (2021). "The Enceladus Orbilander Mission Concept: balancing return and resources in the search for life." *The Planetary Science Journal* 2.2, 77.
- Mahaffy, P. R., Webster, C. R., Atreya, S. K., Franz, H., Wong, M., Conrad, P. G., Harpold, D., Jones, J. J., Leshin, L. a, Manning, H., Owen, T., Pepin, R. O., Squyres, S., Trainer, M., Kempainen, O., Bridges, N., Johnson, J. R., Minitti, M., Cremers, D., ... Moores, J. E. (2013). Abundance and Isotopic Composition of Gases in the Martian Atmosphere from the Curiosity Rover. *Science*, 341(6143), 263–266. <https://doi.org/10.1126/science.1237966>
- Marin-Carbonne, J., Remusat, L., Sforna, M. C., Thomazo, C., Cartigny, P., & Philippot, P. (2018). Sulfur isotope's signal of nanopyrites enclosed in 2.7 Ga stromatolitic organic remains reveal microbial sulfate reduction. *Geobiology*, 16(2), 121–138. <https://doi.org/10.1111/gbi.12275>
- Marshall SM, Murray ARG, Cronin L. (2017). A probabilistic framework for identifying biosignatures using pathway complexity. *Philosophical Transactions of the Royal Society A*, 375: 20160342. <http://dx.doi.org/10.1098/rsta.2016.0342>
- Marty, Bernard, Guillaume Avice, Yuji Sano, Kathrin Altwegg, Hans Balsiger, Myrtha Hässig, Alessandro Morbidelli, Olivier Mousis, and Martin Rubin (2016). "Origins of volatile elements (H, C, N, noble gases) on Earth and Mars in light of recent results from the ROSETTA cometary mission." *Earth and Planetary Science Letters* 441, 91-102.
- McKay, D.S., Gibson, E.K. Jr., Thomas-Keprta, K.L., Vali, H., Romanek, C.S., Clemett, S.J., Chillier, X.D.F., Maechling, C.R. & Zare R.N. (1996). Search for past life on Mars: Possible relic biogenic activity in Martian meteorite ALH84001. *Science*, 273(5277), 924-930.
- McLoughlin, N., Wilson, L. A., & Brasier, M. D. (2008). Growth of synthetic stromatolites and wrinkle structures in the absence of microbes – implications for the early fossil record. *Geobiology*, 6(2), 95–105. <https://doi.org/10.1111/j.1472-4669.2007.00141.x>
- Meadows, V. S. (2017). Reflections on O₂ as a Biosignature in Exoplanetary Atmospheres. *Astrobiology*, 17(10), 1022–1052. <https://doi.org/10.1089/ast.2016.1578>
- Meadows, V. S., Reinhard, C. T., Arney, G. N., Parenteau, M. N., Schwieterman, E. W., Domagal-Goldman, S. D., Lincowski, A. P., Stapelfeldt, K. R., Rauer, H., DasSarma, S., Hegde, S., Narita, N., Deitrick, R., Lustig-Yaeger, J., Lyons, T. W., Siegler, N., & Grenfell, J. L. L. (2018). Exoplanet Biosignatures: Understanding Oxygen as a Biosignature in the Context of Its Environment. *Astrobiology*, 18(6), 630–662. <https://doi.org/10.1089/ast.2017.1727>
- Meixner, M., Cooray, A., Leisawitz, D., Staguahn, J., Armus, L., Battersby, C., Bauer, J., Bergin, E., Bradford, C. M., Ennico-Smith, K., Fortney, J., Kataria, T., Melnick, G., Milam, S., Narayanan, D., Padgett, D., Pontoppidan, K., Pope, A., Roellig, T., ... Webster, C. (2019). Origins Space Telescope Mission Concept Study Report. *ArXiv Preprint 1912.06213*. <http://arxiv.org/abs/1912.06213>.
- Ménez, B., Pisapia, C., Andreani, M., Jamme, F., Vanbellingen, Q.P., Brunelle, A., Richard, L., Dumas, P. & Réfrégiers M. (2018). Abiotic synthesis of amino acids in the recesses of the oceanic lithosphere. *Nature* 564, 59-63.
- Merritt, D.A. & Hayes, J.M. (1994) Factors controlling precision and accuracy in isotope-ratio-monitoring mass spectrometry. *Analytical Chemistry*, 66(14), 2336-2347.
- Meyer, M.A., Kminek, G., Beaty, D.W., Carrier, B.L., Haltigin, T., Hays, L.E., Agee, C., Buseman, H., Cavalazzi, B., Cockell, C.S., Debaille, V., Glavin, D.P., Grady, M.M., Hauber, E., Hutzler, A., Marty, B., McCubbin, F.M., Pratt, L.M., Regberg, A.B., Smith, A.L., Smith, C.L., Summons, R.E., Swindle, T.D., Tait, K.T., Tosca, N.J., Udry, A., Usui, T., Velbel, M.A., Wadhwa, M., Westall, F. & Zorzano, M.-P. (2021). Final report of the Mars Sample Return Planning Group 2 (MSPG2). *Astrobiology*, in press.

- Mustard, J.F., Adler, M., Allwood, A., Bass, D.S., Beaty, D.W., III, J.F.B., Brinckerhoff, W.B., Carr, M., Marais, D.J.D., Drake, B., Edgett, K.S., Eigenbrode, J., Elkins-Tanton, L.T., Grant, J.A., Milkovich, S.M., Ming, D., Moore, C., Murchie, S., Onstott, T.C., Ruff, S.W., Sephton, M.A., Steele, A., Treiman, A., (2013). Report of the Mars 2020 Science Definition Team, 154 pp., posted July, 2013, by the Mars Exploration Program Analysis Group (MEPAG) at http://mepag.jpl.nasa.gov/reports/MEP/Mars_2020_SDT_Report_Final.pdf.
- Nair, H., Allen, M., Anbar, A. D., Yung, Y. L., & Clancy, R. T. (1994). A Photochemical Model of the Martian Atmosphere. *Icarus*, 111(1), 124–150. <https://doi.org/10.1006/icar.1994.1137>
- National Academies of Sciences, E., Medicine (2019). An Astrobiology Strategy for the Search for Life in the Universe. *The National Academies Press*, Washington, DC.
- NASEM (2017) Powering science: NASA's large strategic science missions. *National Academies Press*, Washington, DC. <https://doi.org/10.17226/24857>
- Navarro-González, R., Vargas, E., de la Rosa, J., Raga, A.C. & McKay, C.P. (2010) Reanalysis of the Viking results suggest perchlorate and organics at midlatitudes on Mars. *Journal of Geophysical Research, Planets*, 115, doi:10.1029/2010JE003599.
- Nutman, A. P., Bennett, V. C., Friend, C. R. L., Van Kranendonk, M. J., & Chivas, A. R. (2016). Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature*, 537(7621), 535–538. <https://doi.org/10.1038/nature19355>
- Nutman, A. P., Bennett, V. C., Friend, C. R., Van Kranendonk, M. J., Rothacker, L., & Chivas, A. R. (2019). Cross-examining Earth's oldest stromatolites: Seeing through the effects of heterogeneous deformation, metamorphism and metasomatism affecting Isua (Greenland) ~3700 Ma sedimentary rocks. *Precambrian Research*, 331, 105347. <https://doi.org/10.1016/j.precamres.2019.105347>
- Nutman, A. P., Bennett, V. C., Friend, C. R., & Van Kranendonk, M. J. (2021). In support of rare relict ~3700 Ma stromatolites from Isua (Greenland). *Earth and Planetary Science Letters*, 562, 116850. <https://doi.org/10.1016/j.epsl.2021.116850>
- Olson, S. L., Schwieterman, E. W., Reinhard, C. T., Ridgwell, A., Kane, S. R., Meadows, V. S., & Lyons, T. W. (2018). Atmospheric Seasonality as an Exoplanet Biosignature. *The Astrophysical Journal Letters*, 858(2), L14. <https://doi.org/10.3847/2041-8213/aac171>
- Owen, T. (1980). The Search for Early Forms of Life in Other Planetary Systems: Future Possibilities Afforded by Spectroscopic Techniques. In M. D. Papagianni (Ed.), *Strategies for the Search for Life in the Universe* (pp. 177–185). Springer Netherlands. https://doi.org/10.1007/978-94-009-9115-6_17
- Pappalardo, R. T., D. Senske, H. Korth, R. Klima, S. Vance, and K. Craft (2017). "The Europa Clipper Mission: Exploring the Habitability of a Unique Icy World." In *European Planetary Science Congress*, pp. EPSC2017-304.
- Pasteris, J. D., & Wopenka, B. (2003). Necessary, but Not Sufficient: Raman Identification of Disordered Carbon as a Signature of Ancient Life. *Astrobiology*, 3(4), 727–738. <https://doi.org/10.1089/153110703322736051>
- Petroff, A. P., Sim, M. S., Maslov, A., Krupenin, M., Rothman, D. H., Bosak, T. (2010). Biophysical basis for the geometry of conical stromatolites. *Proceedings of the National Academy of Sciences*.107 (22) 9956-9961. <https://doi.org/10.1073/pnas.1001973107>.
- Pope, M.C., Grotzinger, J.P. * Schreiber, B.C. (2000) Evaporitic subtidal stromatolites produced by in situ precipitation: Textures, facies associations, and temporal significance. *Journal of Sedimentary Research*, 70(5), 1139-1151.
- Postberg, F., Khawaja, N., Abel, B. et al. (2018). Macromolecular organic compounds from the depths of Enceladus. *Nature* 558, 564–568.
- Quanz, S. P., Kammerer, J., Defrère, D., Absil, O., Glauser, A., & Kitzmann, D. (2018). Exoplanet science with a space-based mid-infrared nulling interferometer. In A. Mérand, M. J. Creech-Eakman, & P. G. Tuthill (Eds.), *Optical and Infrared Interferometry and Imaging VI* (Issue July, p. 37). SPIE. <https://doi.org/10.1117/12.2312051>

- Quanz, S. P., Ottiger, M., Fontanet, E., Kammerer, J., Menti, F., Dannert, Gheorghe, F.A., Absil, O., Airapetian, V.S., Alei, E., Allart, R., Angerhausen, D., Blumenthal, S., Buchhave, L.A., Cabrera, J., Carrión-González, Ó., Chauvin, G., Danchi, W.C., Dandumont, C., Defrère, D., Dorn, C., Ehrenreich, D., Ertel, S. Fridlund, M., García Muñoz, A., Gascón, C., Girard, J.H., Glauser, A., Grenfell, J.L., Guidi, G., Hagelberg, J., Helled, R., Ireland, M.J., Kopparapu, R.K., Korth, J., Kozakis, T., Kraus, S., Léger, A., Leedjäv, L., Lichtenberg, T., Lillo-Box, J., Linz, H., Liseau, R., Loicq, J., Mahendra, V., Malbet, F., Mathew, J., Mennesson, B., Meyer, M.R., Mishra, L., Molaverdikhani, K., Noack, L., Oza, A.V., Pallé, E., Parviainen, H., Quirrenbach, A., Rauer, H., Ribas, I., Rice, M., Romagnolo, A., Rugheimer, S., Schwieterman, E.W., Serabyn, E., Sharma, S., Stassun, K.G., Szulágyi, J., Wang, H.S., Wunderlich, F., Wyatt, M.C. & LIFE collaboration. (2021). Large Interferometer For Exoplanets (LIFE): I. Improved exoplanet detection yield estimates for a large mid-infrared space-interferometer mission. arXiv preprint arXiv:2101.07500.
- Rasmussen, B., Fletcher, I.R., Brocks, J.J. & Kilburn, M.R. (2008). Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*, 455(7216), pp.1101-1104.
- Rauer, H., Gebauer, S., Paris, P. v., Cabrera, J., Godolt, M., Grenfell, J. L., Belu, A., Selsis, F., Hedelt, P., & Schreier, F. (2011). Potential biosignatures in super-Earth atmospheres. *Astronomy & Astrophysics*, 529(5), A8. <https://doi.org/10.1051/0004-6361/201014368>
- Rein, H., Fujii, Y., & Spiegel, D. S. (2014). Some inconvenient truths about biosignatures involving two chemical species on Earth-like exoplanets. *Proceedings of the National Academy of Sciences*, 111(19), 6871–6875. <https://doi.org/10.1073/pnas.1401816111>
- Reinhard, C. T., Olson, S. L., Schwieterman, E. W., & Lyons, T. W. (2017). False Negatives for Remote Life Detection on Ocean-Bearing Planets: Lessons from the Early Earth. *Astrobiology*, 17(4), 287–297. <https://doi.org/10.1089/ast.2016.1598>
- Richard, C., Gordon, I. E., Rothman, L. S., Abel, M., Frommhold, L., Gustafsson, M., Hartmann, J. M., Hermans, C., Lafferty, W. J., Orton, G. S., Smith, K. M., & Tran, H. (2012). New section of the HITRAN database: Collision-induced absorption (CIA). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 113(11), 1276–1285. <https://doi.org/10.1016/j.jqsrt.2011.11.004>
- Riding, R. (2011). The Nature of Stromatolites: 3,500 Million Years of History and a Century of Research. In J. Reitner, N.-V. Quéric, & G. Arp, *Advances in Stromatolite Geobiology* (Vol. 131, pp. 29–74). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-10415-2_3
- Robinson, T. D. (2017). Characterizing Exoplanet Habitability. *ArXiv Preprint 1701.05205*. https://doi.org/10.1007/978-3-319-30648-3_67-1
- Robinson, T. D., Ennico, K., Meadows, V. S., Sparks, W., Bussey, D. B. J., Schwieterman, E. W., & Breiner, J. (2014). Detection of Ocean Glint and Ozone Absorption Using LCROSS Earth Observations. *The Astrophysical Journal*, 787, 171. <https://doi.org/10.1088/0004-637X/787/2/171>
- Rosario-Franco, M., Quarles, B., Musielak, Z. E. & Cuntz, M. (2020). Orbital Stability of Exomoons and Submoons with Applications to Kepler 1625b-I. *The Astronomical Journal*, 159(6), 260. <https://doi.org/10.3847/1538-3881/ab89a7>.
- Rouillard, Joti, Mark van Zuilen, Céline Pisapia, and Juan-Manuel Garcia-Ruiz. (2021). "An alternative approach for assessing biogenicity." *Astrobiology* 21, no. 2, 151-164.
- Sagan, C., Thompson, W. R., Carlson, R., Gurnett, D., & Hord, C. (1993). A search for life on Earth from the Galileo spacecraft. *Nature*, 365(6448), 715–721.
- Sarafian, Adam R., Sune G. Nielsen, Horst R. Marschall, Francis M. McCubbin, and Brian D. Monteleone (2014). "Early accretion of water in the inner solar system from a carbonaceous chondrite-like source." *Science* 346, no. 6209, 623-626.
- Schaefer, L., Wordsworth, R. D., Berta-Thompson, Z., & Sasselov, D. (2016). Predictions of the Atmospheric Composition of GJ 1132b. *The Astrophysical Journal*, 829(2), 63. <https://doi.org/10.3847/0004-637X/829/2/63>

- Schopf, J.W. (1993). Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science*, 260(5108), pp.640-646.
- Schopf, J. W. (2006). Fossil evidence of Archaean life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470), 869–885. <https://doi.org/10.1098/rstb.2006.1834>
- Schopf, J. W., Kudryavtsev, A. B., Czaja, A. D., & Tripathi, A. B. (2007). Evidence of Archean life: Stromatolites and microfossils. *Precambrian Research*, 158(3–4), 141–155. <https://doi.org/10.1016/j.precamres.2007.04.009>
- Schwieterman, E. W., Cockell, C. S., & Meadows, V. S. (2015). Nonphotosynthetic Pigments as Potential Biosignatures. *Astrobiology*, 15(5), 341–361. <https://doi.org/10.1089/ast.2014.1178>
- Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., Harman, C. E., DasSarma, S., Fisher, T. M., Arney, G. N., Hartnett, H. E., Reinhard, C. T., Olson, S. L., Meadows, V. S., Cockell, C. S., Walker, S. I., Grenfell, J. L., Hegde, S., Rugheimer, S., Hu, R., & Lyons, T. W. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life. *Astrobiology*, 18(6), 663–708. <https://doi.org/10.1089/ast.2017.1729>
- Schwieterman, E. W., Meadows, V. S., Domagal-Goldman, S. D., Deming, D., Arney, G. N., Luger, R., Harman, C. E., Misra, A., & Barnes, R. (2016). Identifying Planetary Biosignature Impostors: Spectral Features of CO And O4 Resulting From Abiotic O2/O3 Production [Earth and Planetary Astrophysics]. *The Astrophysical Journal*, 819(1), L13. <https://doi.org/10.3847/2041-8205/819/1/L13>
- Schwieterman, E. W., Robinson, T. D., Meadows, V. S., Misra, A., & Domagal-Goldman, S. (2015). Detecting and Constraining N2 Abundances in Planetary Atmospheres Using Collisional Pairs. *The Astrophysical Journal*, 810(1), 57. <https://doi.org/10.1088/0004-637X/810/1/57>
- Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., Harman, C. E., DasSarma, S., Fisher, T. M., Arney, G.N., Hartnett, H. E., Reinhard, C.T., Olson, S.L., Meadows, V.S., Cockell, C. S., Walker, S.I., Grenfell, J.L., Hegde, S., Rugheimer, S., Hu, R., Lyons, T. W. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life, *Astrobiology*, 18, 663.
- Seager, S., Schrenk, M., & Bains, W. (2012). An astrophysical view of Earth-based metabolic biosignature gases. *Astrobiology*, 12(1), 61–82. <https://doi.org/10.1089/ast.2010.0489>
- Seager, S., Turner, E. L., Schafer, J., & Ford, E. B. (2005). Vegetation's Red Edge: A Possible Spectroscopic Biosignature of Extraterrestrial Plants. *Astrobiology*, 5(3), 372–390. <https://doi.org/10.1089/ast.2005.5.372>
- Selsis, F., Despois, D., & Parisot, J.-P. (2002). Signature of life on exoplanets: Can Darwin produce false positive detections? *Astronomy and Astrophysics*, 388(3), 985–1003. <https://doi.org/10.1051/0004-6361:20020527>.
- Sephton, M.A. & Hazen, R.M. (2013) On the origins of deep hydrocarbons. *Reviews in Mineralogy and Geochemistry* 75, 449-465.
- Serindag, D. B. & Snellen, I. A. G. (2019). Testing the Detectability of Extraterrestrial O₂ with the Extremely Large Telescopes Using Real Data with Real Noise. *The Astrophysical Journal*, 871(1), L7. <https://doi.org/10.3847/2041-8213/aafa1f>
- SETI (2018) Declaration of Principles Concerning the Conduct of the Search for Extraterrestrial Intelligence. <https://www.seti.org/protocols-eti-signal-detection>. Retrieved Sept. 2021.
- Schieber, J. & Arnett, H.J. (2003). Nannobacteria as a by-product of enzyme-driven tissue decay. *Geology*, 31, 717-720.
- Sherwood Lollar, B., Westgate, T.D., Ward, J.A., Slater, G.F. & Lacrampe-Couloume, G. (2002). Abiogenic formation of alkanes in the Earth's crust as a minor source for global hydrocarbon reservoirs. *Nature*, 416, 522-524.
- Snellen, I. A. G., de Kok, R. J., le Poole, R., Brogi, M., & Birkby, J. (2013). Finding Extraterrestrial Life Using Ground-based High-dispersion Spectroscopy [Earth and Planetary Astrophysics]. *The Astrophysical Journal*, 764(2), 182. <https://doi.org/10.1088/0004-637X/764/2/182>

- Soutter, A.R.B. and Möttus, R., (2020) "Global warming" versus "climate change": A replication on the association between political self-identification, question wording, and environmental beliefs. *Journal of Environmental Psychology* 69, 101413. <https://doi.org/10.1016/j.jenvp.2020.101413>.
- Sproß, L., Scherf, M., Shematovich, V. I., Bisikalo, D. v., & Lammer, H. (2021). Life as the Only Reason for the Existence of N₂–O₂-Dominated Atmospheres. *Astronomy Reports*, 65(4), 275–296. <https://doi.org/10.1134/S1063772921040077>
- Stocker, T.F., D. Qin, G.-K. Plattner, L.V. Alexander, S.K. Allen, N.L. Bindoff, F.-M. Bréon, J.A. Church, U. Cubasch, S. Emori, P. Forster, P. Friedlingstein, N. Gillett, J.M. Gregory, D.L. Hartmann, E. Jansen, B. Kirtman, R. Knutti, K. Krishna Kumar, P. Lemke, J. Marotzke, V. Masson-Delmotte, G.A. Meehl, I.I. Mokhov, S. Piao, V. Ramaswamy, D. Randall, M. Rhein, M. Rojas, C. Sabine, D. Shindell, L.D. Talley, D.G. Vaughan and S.-P. Xie, (2013). Technical Summary. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Stüeken, E. E., Kipp, M. A., Koehler, M. C., Schwieterman, E. W., Johnson, B., & Buick, R. (2016). Modeling pN₂ through Geological Time: Implications for Planetary Climates and Atmospheric Biosignatures. *Astrobiology*, 16(12), 949–963. <https://doi.org/10.1089/ast.2016.1537>
- Sullivan, Gail M., and Richard Feinn (2012). "Using effect size—or why the P value is not enough." *Journal of graduate medical education* 4, no. 3, 279-282.
- Summons, R.E., Powell, T.G. & Boreham, C.J. (1988). Petroleum geology and geochemistry of the Middle Proterozoic McArthur Basin, Northern Australia: III. Composition of extractable hydrocarbons. *Geochimica et Cosmochimica Acta*, 52(7), 1747-1763.
- Summons, R., Freeman, K.H., Grice, K. & Pancost, R. (2008) Where would we be without the isotopes? *Organic Geochemistry*, 39(5), 483-484.
- Szopa, C., Freissinet, C., Glavin, D.P., Millan, M., Buch, A., Franz, H.B., Summons, R.E., Sumner, D.Y., Sutter, B., Eigenbrode, J.L., Williams, R.H., Navarro-González, R., Guzman, M., Malespin, C., Teinturier, S., Mahaffy, P.R. & Cabane, M. (2020). First detections of dichlorobenze isomers and trichloromethylpropane from organic matter indigenous to Mars mudstone in Gale Crater, Mars: Results from the Sample Analysis at Mars instrument onboard the Curiosity Rover. *Astrobiology*, 20(2), 292-307.
- Tabataba-Vakili, F., Grenfell, J. L., Griebmeier, J.-M., & Rauer, H. (2016). Atmospheric effects of stellar cosmic rays on Earth-like exoplanets orbiting M-dwarfs. *Astronomy & Astrophysics*, 585, A96. <https://doi.org/10.1051/0004-6361/201425602>
- Tessenyi, M., Tinetti, G., Savini, G., & Pascale, E. (2013). Molecular detectability in exoplanetary emission spectra. *Icarus*, 226(2), 1654–1672. <https://doi.org/10.1016/j.icarus.2013.08.022>
- The LUVOIR Team. (2019). The LUVOIR Mission Concept Study Final Report. *ArXiv Preprint 1912.06219*. <http://arxiv.org/abs/1912.06219>
- Tian, F., France, K., Linsky, J. L., Mauas, P. J. D., & Vieytes, M. C. (2014). High stellar FUV/NUV ratio and oxygen contents in the atmospheres of potentially habitable planets. *Earth and Planetary Science Letters*, 385, 22–27. <https://doi.org/10.1016/j.epsl.2013.10.024>
- Turtle, E.P., Trainer, M.G., Barnes, J.W., Lorenz, R.D., Hibbard, K.E., Adams, D.S., Bedini, P., Brinckerhoff, W.B., Cable, M.L., Ernst, C. and Freissinet, C. (2019). March. Dragonfly: In Situ Exploration of Titan's Organic Chemistry and Habitability. In *Lunar and Planetary Science Conference* (No. 2132, p. 2888).
- Van Kranendonk, M. J. (2011). Morphology as an Indicator of Biogenicity for 3.5–3.2 Ga Fossil Stromatolites from the Pilbara Craton, Western Australia. In J. Reitner, N.-V. Quéric, & G. Arp (Eds.), *Advances in Stromatolite Geobiology* (pp. 537–554). Springer. https://doi.org/10.1007/978-3-642-10415-2_32.

- Verrecchia, E.P. (1996) Morphometry of microstromatolites in calcrete laminar crusts and a fractal model of their growth. *Mathematical Geology*, 28, 87-109.
- Villanueva, G.L., Cordiner, M., Irwin, P.G.J., De Pater, I., Butler, B., Gurwell, M., Milam, S.N., Nixon, C.A., Luszcz-Cook, S.H., Wilson, C.F. and Kofman, V. (2021). No evidence of phosphine in the atmosphere of Venus from independent analyses. *Nature Astronomy*, 5(7), pp.631-635.
- Wacey, D. (2010). Stromatolites in the ~3400 Ma Strelley Pool Formation, Western Australia: examining biogenicity from the macro- to the nano-scale. *Astrobiology* 10, 381–395.
- Walker, S. I., Bains, W., Cronin, L., DasSarma, S., Danielache, S., Domagal-Goldman, S., Kacar, B., Kiang, N. Y., Lenardic, A., Reinhard, C. T., Moore, W., Schwieterman, E. W., Shkolnik, E. L., & Smith, H. B. (2018). Exoplanet Biosignatures: Future Directions. *Astrobiology*, 18(6), 779–824. <https://doi.org/10.1089/ast.2017.1738>
- Westall, F., de Ronde, C.E.J., Southam, G., Grassineau, N., Colas, M., Cockell, C. & Lammer H. (2006) Implication of a 2.472-3.333 Gyr-old subaerial microbial mat from the Barberton greenstone belt, South Africa for the UV environmental conditions on the Early Earth. *Philosophical Transactions of the Royal Society B*, 361(1747), <https://doi.org/10.1098/rstb.2006.1896>
- Westall, F., Hickman-Lewis, K., Hinman, N., Gautret, P., Campbell, K.A., Bréhéret, J.G., Foucher, F., Hubert, Az, Sorieul, S., Dass, A.V., Kee, T.P., Georgelin, G. & Brack, A. (2018) A hydrothermal-sedimentary context for the origin of life. *Astrobiology*, 18(3), 259-293.
- Whitmarsh, L. (2009) What's in a name? Commonalities and differences in public understanding of “climate change” and “global warming”. *Public Understanding of Science* 18(4), 401-420. <https://doi.org/10.1177%2F0963662506073088>.
- Williford, K.H., Farley, K.A., Stack, K.M., Allwood, A.C., Beaty, D., Beegle, L.W., Bhartia, R., Brown, A.J., de la Torre Juarez, M., Hamran, S.E. and Hecht, M.H. (2018). The NASA Mars 2020 rover mission and the search for extraterrestrial life. In: *From habitability to life on Mars* (pp. 275-308). Elsevier.
- Wordsworth, R. D., Schaefer, L. K., & Fischer, R. A. (2018). Redox Evolution via Gravitational Differentiation on Low-mass Planets: Implications for Abiotic Oxygen, Water Loss, and Habitability. *The Astronomical Journal*, 155(5), 195. <https://doi.org/10.3847/1538-3881/aab608>
- Wordsworth, R., & Pierrehumbert, R. (2014). Abiotic Oxygen-dominated Atmospheres On Terrestrial Habitable Zone Planets [Earth and Planetary Astrophysics]. *The Astrophysical Journal*, 785(2), L20. <https://doi.org/10.1088/2041-8205/785/2/L20>
- Zahnle, K., Haberle, R. M., Catling, D. C., & Kasting, J. F. (2008). Photochemical instability of the ancient Martian atmosphere. *Journal of Geophysical Research*, 113(E11), E11004. <https://doi.org/10.1029/2008JE003160>
- Zawaski, M. J., Kelly, N. M., Orlandini, O. F., Nichols, C. I. O., Allwood, A. C., & Mojzsis, S. J. (2020). Reappraisal of purported ca. 3.7 Ga stromatolites from the Isua Supracrustal Belt (West Greenland) from detailed chemical and structural analysis. *Earth and Planetary Science Letters*, 545, 116409. <https://doi.org/10.1016/j.epsl.2020.116409>
- Zawaski, M. J., Kelly, N. M., Orlandini, O. F., Nichols, C. I., Allwood, A. C., & Mojzsis, S. J. (2021). Reply: The Isua (Greenland) “relict stromatolites” cannot be confidently interpreted as original sedimentary structures. *Earth and Planetary Science Letters*, 562, 116851. <https://doi.org/10.1016/j.epsl.2021.116851>
- van Zuilen, M.A., Lepland, A., & Arrhenius G. (2002). Reassessing the evidence for the earliest traces of life. *Nature*, 418, 627-630.