Mineralogical Magazine (2024), **88**, 421–429
doi:10.1180/mgm.2024.19

Article

Abundances, distributions and patterns of discovery of new minerals

Carl N. Drummond
Earth and Planetary Science, Department of Physics, Purdue University Fort Wayne, Fort Wayne, Indiana, USA

Abstract

Mineral species are known to be heterogeneously distributed throughout the Earth such that a relatively small number of minerals make up a large proportion of the lithosphere while the majority of all known minerals are rare and have been identified at only a small number of locations that frequently exhibit high levels of species richness. Intuitive understandings of mineral scarcity and abundance are reconsidered through the characterisation of the quantitative aspects of spatio-temporal trends in new mineral discovery. Using data drawn from online mineralogical databases, it is found that the Earth's mineral hotspots exhibit an exponential distribution of species abundance, while those same mineral hotspots exhibit a power-law distribution in the number of minerals first recognised at those locations. That is, locations rich in first occurrences are extremely rare, even when considering only the Earth's most species-rich mineral locations. Global distributions of mineral scarcity and abundance can be estimated from the number of mineral-location pairs for each species reported in a database. Two-thirds of all known species have been reported from ten or fewer locations and the frequency distribution of these mineral-location pairs exhibit a power-law distribution that extends with increasing dispersion over several orders of magnitude of mineral abundance. Initially, nearly all minerals are first reported from only their type locality. Over time, additional occurrences of newly discovered minerals are reported at an average rate of one new location per mineral every 5.5 years. As a result, the percentage of minerals that were discovered in a given year that continue to be known only from their type locality is found to decline exponentially over time. However, a few minerals remain known from only their type locality for long periods, including some that were first identified in the 19th Century. Conversely, other recently identified minerals have been subsequently recognised at locations spanning a wide geographic range such that the number of minerals with cosmopolitan distributions is found to increase exponentially over time. Taken together, these several quantitative representations of mineral distributions lend structure and refinement to qualitative and intuitive notions of the scarcity and abundance of Earth's many minerals.

Keywords: type localities; mineral-location pairs; power-law distributions; endemic minerals; cosmopolitan minerals

(Received 11 July 2023; accepted 13 March 2024; Accepted Manuscript published online: 21 March 2024; Associate Editor: Edward Sturgis Grew)

Introduction

Mineral species are recognised to have a heterogeneous distribution throughout the Earth. Most of the known minerals were formed and reside at the Earth's surface or within the outermost portion of the Earth's lithosphere. Because the most common mineral forming processes recur broadly across space and time (crystallisation from igneous melts, from concentrated saline solutions, from weathering and redox reactions at or near the Earth's surface, or through recrystallisation at elevated temperatures and pressures, often in the presence of reactive fluids) and because a relatively small number of elements are typically found in abundance at these various locations (Christy, 2015), a large proportion of the Earth's lithosphere is composed of a relatively limited number of mineral species (Hazen and Ausubel, 2016). Conversely, a wide range of much more uncommon processes of mineral formation are known to be highly localised in space and time. Additionally, these uncommon processes often have the tendency to concentrate, physically and chemically, relatively rare elements in ways that result in the formation of minerals with uncommon or even unique chemistries. Because these rare elements are often chemically incompatible in common crystallographic structures, their presence in abundance due to localised and perhaps ephemeral mineral forming processes can also result in the formation of uncommon or even unique crystal structures. As such, the nearly 6000 mineral species currently recognised by the Commission on New Minerals, Nomenclature and Classification (CNMNC) of the International Mineralogical Association (IMA, Pasero, 2023) consist of a small number of very commonly occurring minerals and a large number of very rare minerals. Herein these concepts of scarcity and abundance are further elucidated by exploring the concept of mineral diversity hotspots, the scaling relationships present in the abundances of mineral species found at these locations, and the varying gradation from endemic to cosmopolitan distributions of recently discovered mineral species. Taken together, these observations provide the foundation for a more robust quantitative understanding of the temporal and spatial trends in the relative scarcity and abundance of mineral species globally.

Email: drummond@pfw.edu

Cite this article: Drummond C.N. (2024) Abundances, distributions and patterns of discovery of new minerals. *Mineralogical Magazine* **88**, 421–429. https://doi.org/10.1180/mgm.2024.19

The recognition of a new mineral species occurs as the result of the convergence of a variety of factors. First, conditions must have existed at some point in Earth history for the formation of a particular mineral species. Second, irrespective of the mode of mineral formation, some combination of natural (tectonic, volcanic, erosional) or exploratory (excavation, quarrying, mining) processes must have occurred to bring the previously unknown mineral to a place where it becomes available to be discovered. Third, scientists must be motived to undertake the effort necessary to recognise, characterise, and ultimately report the discovery. These three factors, acting in concert, have resulted in the focusing of new mineral discovery to a relatively small number of geographically isolated locations characterised by great mineralogical diversity that can be considered geological analogues of the widely discussed concept of biodiversity hotspots (Myers, 1988; Reid, 1998; Marchese, 2015). Herein, these mineralogical hotspots are defined on the basis of the number of mineral species that have been identified at that locality (observed diversity) as well as by the number of species that were first recognised at that locality (proto-diversity), such that those locations are designated as the type locality of one or more mineral species (Back and Mandarino, 1999; Atencio, 2000). Based on these criteria a set of geologically and geographically unique locations rich in mineral diversity, while also being the sites of a significant number of first discoveries of previously unknown species, are considered in the following analysis.

This study utilises mineral data drawn from two open-access mineralogical databases. The primary and most definitive source of data is the list of minerals approved by the IMA–CNMNC (Downs, 2006; rruff.info/ima; http://cnmnc.units.it/master_list/IMA_Master_List_(2024-03).pdf). Likewise, for information regarding locations of minerals – both the total number of reported locations of specific minerals as well as lists of all minerals known from specific locations including type locality designations – the database hosted by the Hudson Institute of Mineralogy is utilised: 'Mindat' (https://www.mindat.org/). Previous studies have made good use of these databases (Grew and Hazen 2014; Hazen *et al.*, 2015, Hystad *et al.*, 2015b; Grew *et al.*, 2017; Bermanec *et al.*, 2022). To supplement the data obtained from these online sources, selections from the primary literature have also been consulted.

Hotspots of mineral diversity

The locations on Earth characterised by a large number of mineral species, herein known as mineral diversity hotspots, are defined as a single location, or in some cases a small number of sub-localities within a geographically limited area, typically associated with a single or small number of processes of mineral formation. Commonly located within mining districts or in areas of volcanic activity, some of these hotspots were known in antiquity (Mount Vesuvius, Naples, Italy: Pelloux, 1927; Russo *et al.*, 2014) while others have only become known or have formed in recent decades (Tolbachik volcano, Kamchatka, Russia: Fedotov and Markhinin, 1983). Many are associated with fairly well understood geological processes such as volatile-rich pegmatite crystallisation (Hagendorf Pegmatite, Bavaria, Germany: Mücke, 1981) or porphyry copper hydrothermal mineralisation (Chuquicamata Mine, Antofagasta, Chile: Bandy, 1938; Cook, 1978) whereas others are associated with rare events of extraterrestrial origin (Allende meteorite, Chihuahua, Mexico: Fuchs, 1971; El Goresy, *et al.*, 1977). Some are very well known to students of geology (Franklin and Sterling Hill mines New Jersey, USA: Dunn, 1995; Sapucaia mine, Minas Gerais Brazil: Cassedanne and Baptista, 1999; Baijot *et al.*, 2012) whereas others are as yet far less celebrated (Halamish wadi, Hatrurim Basin, Israel: Britvin *et al.*, 2022). In all cases, however, these hotspots are characterised by the occurrence of a large number of minerals, and critical for this study, an abundance of previously unknown mineral species defined by type location designations.

A group of localities with a large number of known mineral species has been drawn from a compilation of such locations identified in the Mindat database. Distributed across all continents excepting Antarctica, these locations exhibit an observed total diversity of over 10,000 mineral-locality pairs (Hystad *et al.*, 2015a, 2015b), with an average of 113 recognised mineral species per location. The Clara mine, Baden-Wurttemberg, Germany (Markl *et al.*, 2019) is currently recognised as the location with the greatest mineral species abundance, with over 460 mineral species identified. Additionally, all of the locations considered are the type locality of at least eight previously unknown mineral species, a cutoff that originates from the Mindat compilation. The location with the largest number of endemic mineral species in the database is the Tolbachik volcano, Kamchatka, Russia, the type locality of more than 140 mineral species. This location is composed of several cinder cone and fumarole sub-localities associated with multiple episodes of recent volcanic activity. Taken together, these hotspots are the type localities of over 1500 minerals, more than a quarter of all currently known mineral species. As such, they provide a robust representation of global patterns in the characteristics of clustering exhibited by mineral species. Importantly, it is recognised that when working in the field, mineralogists and well-trained amateurs do not record or report every occurrence of every mineral they might observe. The strengths, challenges and biases inherent in the use of online databases have been fully discussed elsewhere (Hazen *et al.*, 2015; Alroy *et al.*, 2001; Alroy *et al.*, 2008).

In order to begin the exploration of the quantitative characteristics of the Earth's mineral hotspots, it is appropriate to consider the frequency distributions of the number of mineral species identified (observed diversity) as well as the number of type locality designations (proto-diversity) recorded for each such location. In geological systems, it is often useful to evaluate highly skewed frequency distributions by rank-ordering the magnitudes of all observations in order to obtain a function $E(f)$ where E is the number of locations with more occurrences than the value f. The shape of graphs of this function provide insights into the statistical characteristics of the observed distribution, as well as suggesting potentially significant causal factors that might underly the formation of such distributions (Turcotte, 1994; Rothman *et al.*, 1994; Drummond, 1999). This rank-ordering technique, known as exceedance analysis, has been widely utilised in the natural sciences because it eliminates the necessity of creating data bin sizes of arbitrary magnitude. Exceedance plots of the observed and proto-diversity present within the group of hotspots considered herein are characterised by strongly concave-up curves (Fig. 1a and c). However, logarithmic transformations of the exceedance values illustrate significant differences between these two distributions. In the case of the observed diversity distribution, the base 10 logarithm of the exceedance values demonstrates that these mineral species-richness data define a log-linear or exponential function such that:

$$\log_{10}E(f) = (-4.44\times10^{-3})f + 2.06\,(r^2 = 0.9916)$$

That is, mineral species richness at these hotspots is exponentially distributed such that a small number of locations possess very large amounts of mineral diversity (lower right Fig. 1a and b)

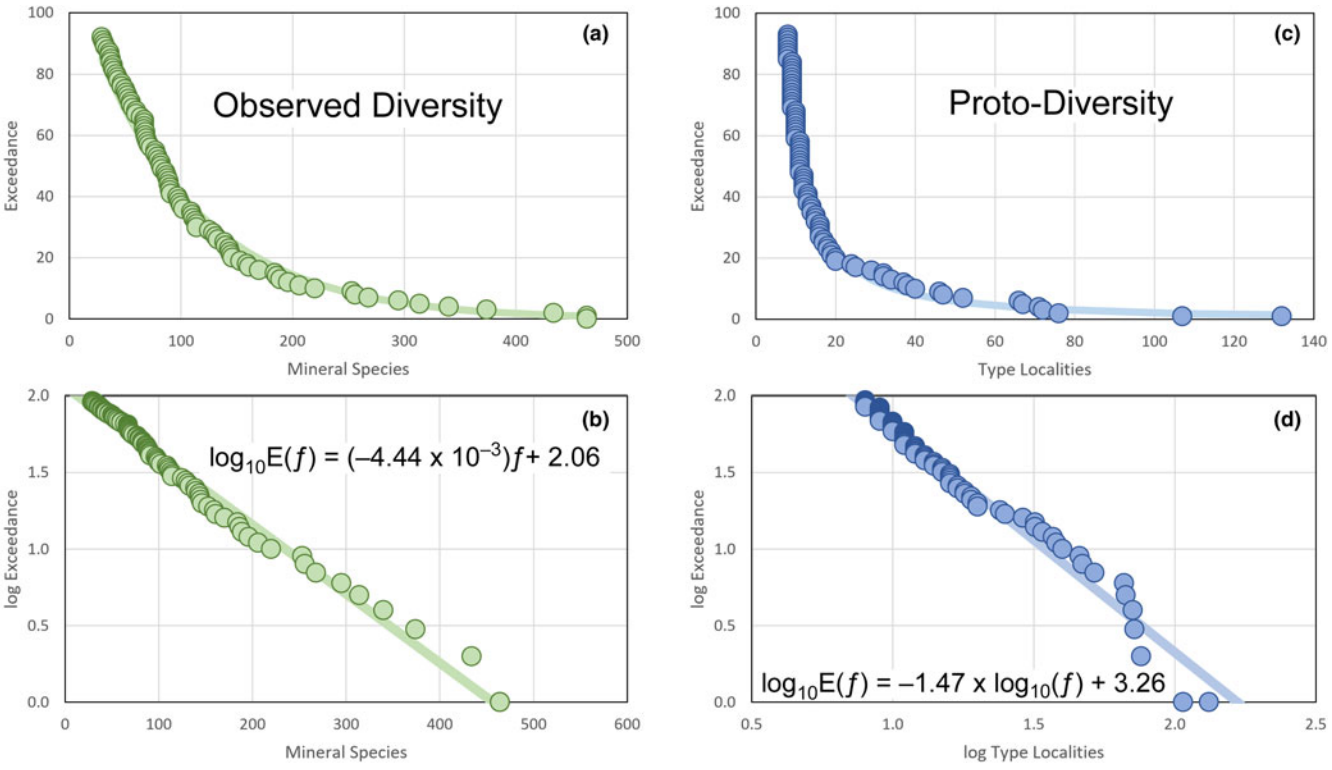

Figure 1. Mineral hotspots are defined as those locations, or groups of sublocations clustered in close geographic proximity, that are characterised by either an abundance of mineral species (observed diversity) or an abundance of minerals first discovered at that location (proto-diversity). Using the formulation E(*f*) where E is the number of locations with more minerals than the value *f*, known as the exceedance value, data are displayed from the most mineral-rich hotspots (Mindat.org). (a) The observed diversity exceedance distribution follows an exponential curve; (b) [$\log_{10}$ E(*f*) = (–4.44 × 10^3)*f* + 2.06]. The concavity of the proto-diversity curve (c) is much greater than the observed diversity data (a) such that the proto-diversity data define a power-law exceedance relationship (d) [$\log_{10}$ E(*f*) = –1.47 × $\log_{10}$(*f*) + 3.26] highlighting the extreme scarcity of locations rich in previously unknown minerals.

while most other locations possess far fewer total minerals (upper left Fig. 1a and b); and there exists a mathematical relationship that describes the distribution of mineral richness present within these locations (Fig. 1b).

The curvature of the first occurrence proto-diversity data is much greater than that found for the observed diversity data (Fig. 1c *vs*. 1a). For this case, base 10 logarithms of both the exceedance values and the number of type locality designations result in the linearisation of the distribution to give:

$$\log_{10}\mathrm{E}(f) = -1.47 \times \log_{10}(f) + 3.26 \quad (r^2 = 0.9889)$$

From this, it is clear that even among the most locations with many known mineral species, those with high levels of proto-diversity are rare (Fig. 1d), such that those locations that have contributed the most to the ongoing expansion in the number of known minerals are orders of magnitude less common than localities that have produced a significant but more modest number of new minerals, thereby highlighting the importance of the mineral hotspot concept to the history and practice of mineral discovery, particularly in the latter half of the 20th and first decades of the 21st century. Such highly curved power-law distributions which are linearised by taking logarithms of both variables have been identified in a variety of complex and often self-organising natural systems (Corral and González, 2019, for a comprehensive summary). The potential origins of these two different types of curves begin to become appreciable when one further considers the quantitative meaning of the distributions in the context of mineral forming processes. Interestingly, Hummer (2021) has also identified a fractal like power law relationship in the number of mineral species populating the various point groups of the crystal systems.

As noted above, most mineral forming processes result in the crystallisation of a relatively small number of common minerals under conditions that frequently recur in time and space, such that at most locations within the Earth's outer lithosphere the observed mineral diversity is low. Previous studies have considered the spatio-temporal relationships in new mineral discovery globally (Ponomar *et al*., 2023) as well as the clustering of chemical compositions using network analysis of rarity groups (Gavryliv *et al*., 2022). As such, locations that are characterised by a very large number of mineral species must have been subjected to mineral forming processes that were characterised by chemical and physical conditions that were highly uncommon (Hazen and Ausubel, 2016). One possibility is that the greater the number of minerals observed at a location the more uncommon, perhaps the more extreme, the conditions must have been at the time of mineral formation. Additionally, one could consider the possibility that many of the locations with high levels of observed mineral diversity might have been subject to multiple phases of mineral formation under evolving conditions of temperature, pressure, or volatile content and chemical composition. Beyond efforts to identify new mineral species, mineralogists and petrologists work to tease out the specifics of the types and styles of paragenesis recorded at these geologically complex locations and it is information about the various modes of mineral formation that fully informs our understanding of mineral species abundance and scarcity relationships.

The statistical characteristics present within the frequency distribution of proto-diversity of hotspot locations suggests the identification of previously unknown minerals can be distinctly different from the identification of other rare but previously known species. As noted above, identification of a new mineral requires the convergence of a complex combination of geological processes and human activities, first for the formation of the mineral and then for its recognition. The complexity of the interrelationships of these processes is best highlighted by the histories of discovery of new minerals at hotspots. In some cases, new minerals have been discovered at a single locality at a fairly steady rate for up to 250 years, while at others new mineral discovery has accelerated rapidly due to a combination of geological and observational causes over a matter of only a few decades (Fig. 2). Further, the discovery history of unknown minerals is driven by complex interrelationships between several factors. First, how many previously unknown mineral species are present and available for discovery? Second, how rare are the as yet undiscovered minerals? Crystal-chemical structures that exist as only a few unit cells clustered together along the margins of previously known mineral grains are unlikely to ever be recognised, by even the most careful and technically advanced analytical techniques (Caraballo *et al.*, 2015; Grew *et al.*, 2017). Finally, how do advances in technology (Grey, 2022) and focused human effort impact on the search for unknown minerals? It is likely that human factors play a significant role in determining the history of mineral discovery in that once a site becomes recognised as having a significant amount of endemic mineral diversity, efforts of individuals and research teams directed at searching for, analysing, characterising and reporting new mineral occurrences tend naturally to become concentrated at that location. An example of this phenomenon can be observed in the discovery history of new minerals identified at the Yadovitaya fumarole, second scoria cone, Northern Breakthrough sublocality of the Great Fissure eruption of the Tolbachik volcano, Kamchatka, Russia, where more than 20 new minerals have been discovered in an area of only 2 square metres (Pekov *et al.*, 2020).

Finally, it is worthy of noting that the proto-diversity present at a given location is a subset of the observed diversity for that location such that in most locations there will be many more previously known than newly discovered minerals. The majority of mineral hotspots are found to have proto-diversity values that range between 10% and 30% of the documented observed diversity at that location (Fig. 3). The difference in curve shape is a consequence and demonstration of the fact that more favourable factors need to converge for a new species to be discovered than for a previously known species to be found at a new locality.

Scaling relationships in the abundance of known mineral species

In the following, consideration is shifted from the analysis of mineral diversity distributions at specific locations to the analysis of the global distribution of all known mineral species. As has been noted, it is well understood that there are a large number of rare minerals and there are a small number of common minerals (Hazen and Ausubel, 2016). How can these concepts of relative mineral scarcity and abundance be better quantified and understood? The database Mindat.org, hosted by the Hudson Institute of Mineralogy (Mindat, 2023) provides chemical, crystallographic and taxonomic data on the nearly 6000 minerals approved by the IMA. The database also tabulates the occurrence of these species globally in the form of over 1.4 million mineral-locality pairing reports (Hystad *et al*, 2015a, 2015b). Herein these mineral-locality pairs are used to explore quantitative relationships in the scarcity and abundance of minerals. Importantly, what follows is not an effort to estimate the total number of minerals that exist on Earth (Skinner and Skinner, 1980, Hystad *et al.*, 2015a). Rather, herein a scaling relationship in mineral occurrence data is recognised and characterised in order to shed light on another facet of the continuum of abundances and scarcity exhibited by Earth's minerals.

As described in the preceding discussion of mineral hotspots, the identification of minerals, be they well-known or previously undiscovered, requires the somewhat fortuitous convergence of geological processes and human actions. As not all minerals that occur on Earth have been identified yet (Hystad *et al.*, 2015a; Hazen and Ausubel, 2016; Hazen *et al.*, 2022), and because not all locations of mineral occurrences have been, or perhaps even could be, fully catalogued (Hystad *et al.*, 2015b), any compilation of the occurrences of minerals must be an underrepresentation of total global mineral diversity. However, the Mindat database is a constantly growing and evolving archive of known minerals, their geological context and geographic occurrence. Though certainly incomplete, Mindat has grown to a sufficient size and scope as to be taken as a reasonably representative summary of broad trends in global mineral distribution. The following evaluates the degree to which the mineral-location pair data proves helpful in understanding trends in the global distribution of mineral diversity.

If a mineral is known to occur only at its initial location of discovery, it is said to be endemic to that site. More than two-thirds of the nearly 6000 mineral species have been identified at 10 or fewer locations, and more than a quarter of the 6000 minerals are known only from a single location (Fig. 4a). The highly endemic nature of the majority of known minerals stands in contrast to the global abundance of a small number of common minerals such that less than 10% of all minerals are known to occur at more than 100 documented locations and less than 3% have been reported from more than 1000 locations; these general abundance relationships are well known both intuitively and qualitatively. When graphing the number of mineral species known to occur at a given number of locations a negative log–log linear relationship is observed:

$$\log_{10}\mathrm{F}(\lambda) = -1.20\log_{10}\lambda + 3.19 \quad (r^2 = 0.9975)$$

Where $\mathrm{F}(\lambda)$ is a function that describes the frequency of mineral species known to occur at the number of locations λ. This linear relationship is based on a regression of mineral-location data drawn from those minerals known to occur at ten or fewer locations (Fig. 4a), a subset that accounts for more than two-thirds of all known species. Importantly however, this linear relationship is found to also extend over several orders of magnitude of decreasing mineral frequency and increasing mineral-location pair abundance (Fig. 4b). It has been postulated that mineral species diversity could be fractal (Grew *et al.*, 2017), such that closer study at microscopic scale would result in the identification of many more currently unrecognised species. The presence of these power-law, or fractal, relationships within mineral abundance strongly suggests that it is highly likely the potential for

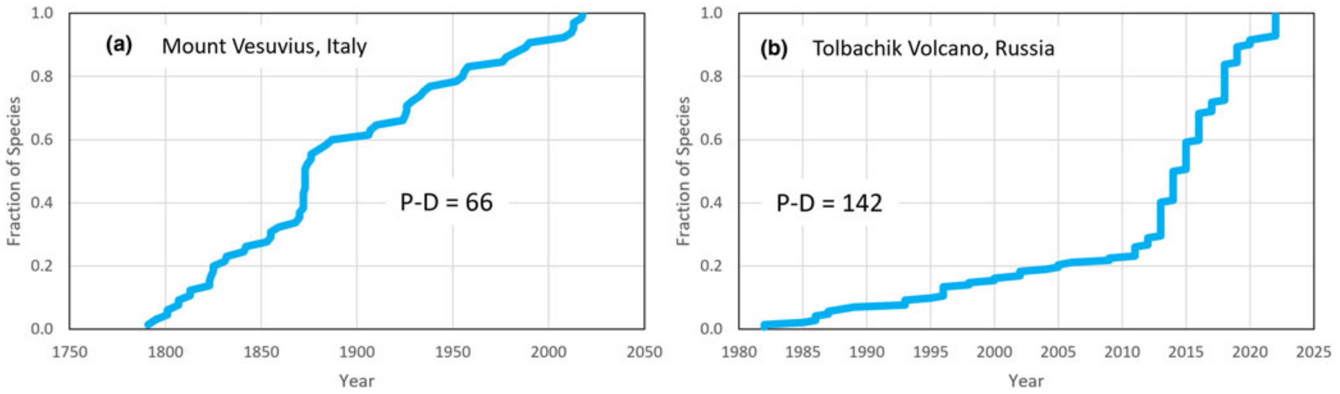

Figure 2. Mineral discovery curves from two hotspots illustrating different temporal trends in discovery. Proto diversity (P-D) is defined as the number of minerals for which each location is the type locality. (a) The curve for Mount Vesuvius displays a nearly linear increase in the number of new mineral discoveries over a period of over 200 years. (b) Conversely, the Tolbachik Volcano locality is characterised by a significant increase in the rate of mineral discovery after the great fissure eruption of 2012–2013. In both graphs the vertical axis is the fraction of currently known mineral species that had been identified at a given time.

discovery of new minerals will not be exhausted in the near future. The Earth's total mineral inventory has been estimated previously by the 'large number of rare events' model (Hystad *et al.*, 2015a).

Dispersion around this power-law trend increases as the number of locations considered expands (Fig. 4b). While this might merely be an inevitable result given the arbitrariness of a non-systematic process of observing and reporting mineral-location pairs, it is sufficiently striking to be worthy of further consideration. The highly skewed nature of the dispersion of mineral-location pair data is most clearly illustrated by reviewing the distribution of location values linked to a single mineral. That is, the set of minerals exhibiting a unique abundance of observed locations (Fig. 5) – recognised as those datapoints falling along the abscissa axis of the graph in Fig. 4b ($\log_{10}$ frequency = 0). The frequency distribution of this representational set of commonly occurring minerals is found to be highly skewed even after taking the logarithm of location frequency (Fig. 5). The power-law relationship calculated from the data in Fig. 4a intersects the abscissa axis within one bin of the modal value of the distribution of this population, thus suggesting that the underlying relationships that govern the large number of rare minerals maintains significant influence in controlling the distribution of common minerals.

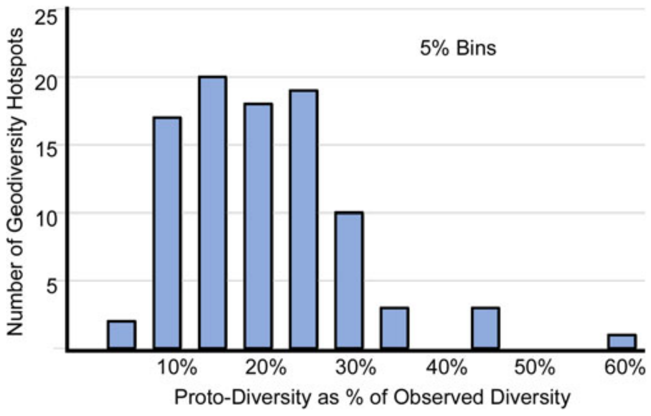

Figure 3. The majority of the mineral diversity hotspots considered have proto-diversity values that range between 10% and 30% of the observed diversity for that location, illustrating the rarity of new mineral discoveries relative to the re-identification of previously known minerals, even in the most mineral-rich settings on Earth.

Further, those minerals with location data most frequently reported in the Mindat database share some interesting characteristics (Table 1). When considering the ten most frequently reported minerals in the database, six were known to the ancients (effective discovery date of 0), seven are either precious metals or ore minerals, three are minerals commonly associated with the products of chemical weathering of rocks at the Earth's surface or as refractory minerals concentrated by weathering, and two are common minerals found in the continentally abundant rock type granite. Consideration of these most widely recognised species in the Mindat database provides more than just a summary of the extreme tail of the mineral occurrence frequency distribution (Fig. 4b). Rather, from these observations it is possible that several biases could impact the reporting of mineral-location data of common minerals that do not, by and large, impact the mineral-location relationship exhibited by those less common minerals (Fig. 4a). Such biases could act to cause the observed increasing dispersion in the number of reported mineral-location pairs and the volumetric abundance of common minerals. The postulated biases in reporting common mineral locations can take several forms. First the number of location reports can be greater than the proportional abundance (volumetric or geographic) of a mineral species. An overreporting bias of this type exists for precious metals, primary ore minerals, and minerals of unique interest to collectors. In such cases a mineral with a disproportionately high number of reported mineral-location pairs would be shifted to the right of what might be its expected abundance on plots of location frequency distributions given a globally random sampling of mineral occurrences (Fig. 4b). Alternatively, commonly occurring but economically and scientifically less-compelling minerals could be underrepresented in the mineral-location pair data due to a bias against re-reporting these widely distributed minerals, which results in their being shifted to the left on plots of location-frequency distributions. These two biases are probably responsible for a significant fraction of the observed dispersion for minerals with more than 100 reported locations (Fig. 4b).

The Earth's most common minerals, however incompletely defined by mineral-location data, must be the product of the co-location of the Earth's most common elements with the most frequently occurring modes of mineral formation (Hazen

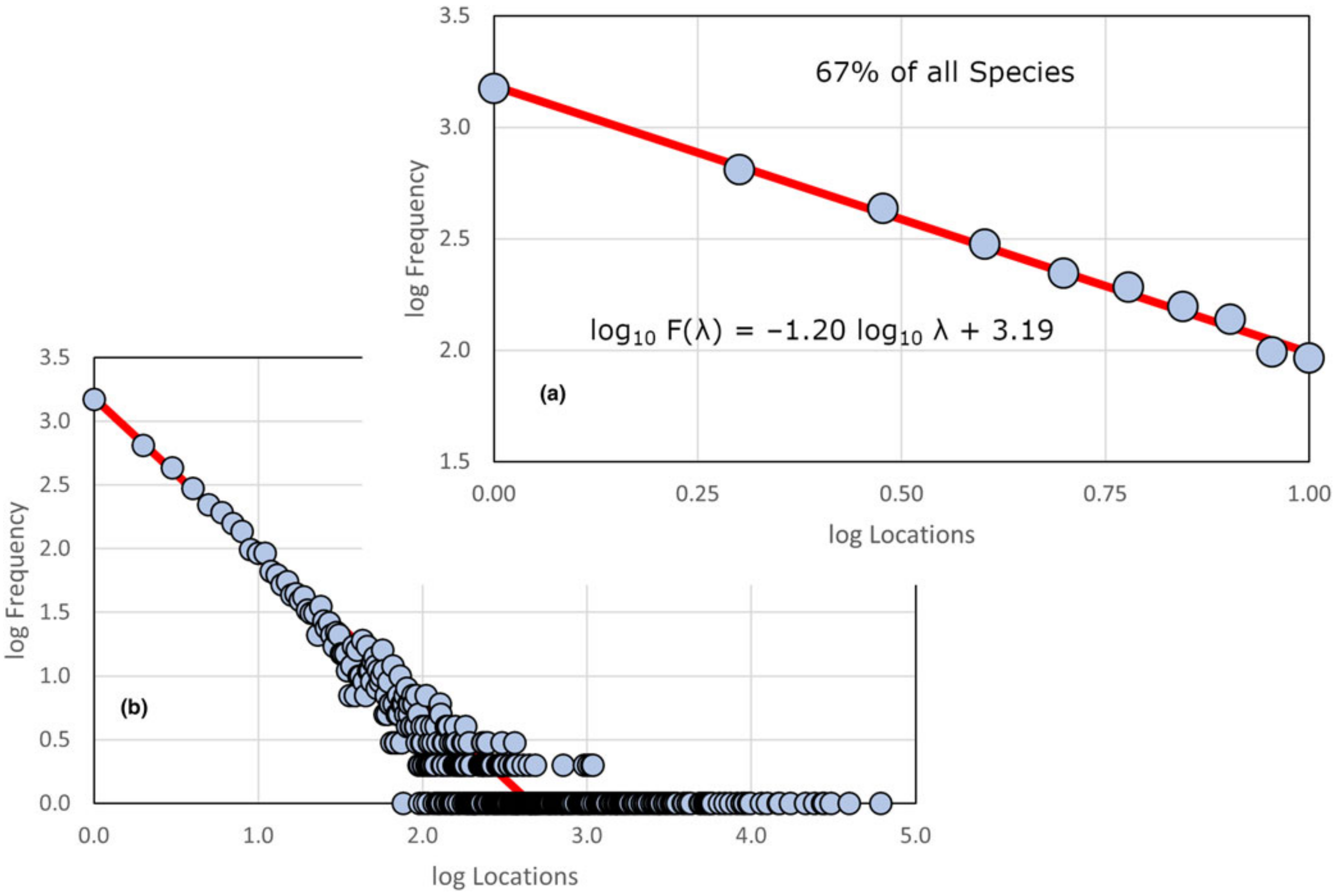

Figure 4. (a) Utilising the community-sourced database Mindat.org to harvest mineral-location pairs for all IMA accepted species in January, 2021, a power-law relationship between the number of species and the number of locations from which they are known is observed with minerals known from 10 or less locations [$\log_{10} F(\lambda) = -1.20 \log_{10} \lambda + 3.19$] which accounts for more than two thirds of all known species. (b) This power-law relationship is found to extend over several orders of magnitude with increasing dispersion around the regression as the logarithm of location frequency increases.

et al., 2015). However, there is an important distinction to be made between those minerals that have a large number of discreet and potentially highly valuable occurrences such as gold and those that make up a significant volume of geological material exposed at or near the Earth's surface, such as feldspar.

From these various considerations it is reasonable to conclude that the analysis of mineral location frequency data is most applicable, without being concerned by community sourced database biases, to those minerals that are rare in occurrence but that also comprise the majority of all known species. Likewise, such an approach is increasingly less applicable to the long tail of the location frequency distribution consisting of a smaller number of minerals with large and increasingly variable location frequencies.

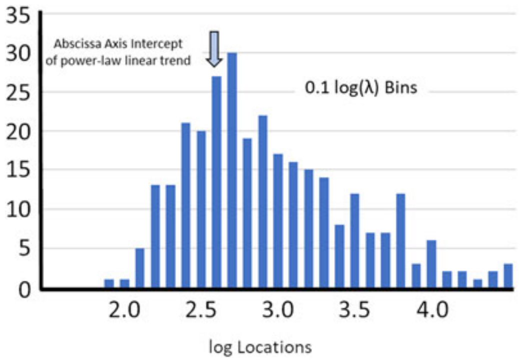

Figure 5. Frequency distribution of those mineral-location pair data that plot along the abscissa axis of Fig. 4b in 0.1 $\log_{10}(\lambda)$ bins, illustrating the highly skewed nature of the dispersion of commonly occurring minerals. Importantly, the power-law trend calculated in in Fig. 4a intersects the abscissa axis within one bin of the modal value of this distribution.

Table 1. Ten mineral species with the most abundantly reported mineral-location data in the Mindat community-sourced database. Six species were known in antiquity (date of 0), seven are metals or ores, three are associated with chemical weathering at the Earth's surface, and two are common minerals found within the most abundant rock type within the continental crust, granite.

Mineral	Date	Locations	Type
Quartz	0	61,156	chemical weathering, granite
Pyrite	0	39,462	metals or ores
Gold	0	30,554	metals or ores
Calcite	0	27,770	chemical weathering
Chalcopyrite	1725	27,198	metals or ores
Galena	0	24,243	metals or ores
Sphalerite	1847	21,482	metals or ores
Muscovite	1850	17,380	granite
Magnetite	1789	14,899	metals or ores
Hematite	0	14,640	chemical weathering

Endemic and cosmopolitan distributions of recently discovered minerals

In addition to recording the geographic distribution of mineral occurrences, the Mindat database includes the date of acceptance of new mineral species by the IMA–CNMNC along with a wealth of information on literature citations of the species, which frequently detail a complex history of understanding, or in some cases misunderstanding and misidentification, of the accepted mineral prior to or after its acceptance. From these data, it is possible to quantify the number of newly recognised minerals approved by the CNMNC in a given year. Because the identification of a new mineral is based upon its recorded occurrence at a specific location – its type locality – in the year of its acceptance, it is expected that the vast majority of all minerals will initially only be known from a single type locality. Rare exceptions to this can occur when minerals are reclassified either due to analytical or technological advances or due to the reassessment of complex mineral series (Leake, 1978; Burke and Leake, 2005; Hawthorne *et al.*, 2012). The assignment of a date has become a formal and official process since the creation of the CNMNC (previously known as the Commission on the Minerals and Mineral Names, CNMMN) in 1959, whereas for all minerals 'grandfathered' into the IMA list of accepted minerals prior to that date the official 'age' of the mineral species is derived from literature sources.

In many cases, over time, additional locations for the occurrence of known minerals will be documented and added to online databases. Some minerals are ultimately recognised from a large and growing number of locations and can therefore be considered to have a cosmopolitan global distribution, while others remain known from only a small number of locations or are even exclusively endemic to their type locality. The process of ongoing documentation of additional mineral occurrences can best be understood by looking back to the number of new occurrences of minerals identified in previous years. That is, it is possible to use the Mindat database as 'time machine' to explore the history of discovery of additional mineral localities of recently discovered minerals. In order to do so, the frequency of occurrence of minerals accepted by the IMA in the years 1990 through 2020 have been evaluated. Due to the database's ongoing growth and evolution, occurrence data for minerals first discovered in those years was harvested from the database on January 19th, 2021 thereby creating a temporal snapshot of knowledge about the distribution of newly discovered minerals. At that time, all of the 93 minerals accepted during the 2020 year were known from only a single location. Looking further back in time, the percentage of minerals added in previous years that remained endemic to a single location declines in a systematic way (Fig. 6a), such that only 9 of the 47 minerals added in 1990 continue to be known from only a single location in 2020. That is, over the 30 years since their acceptance in 1990, 38 minerals were found in one or more locations beyond their type locality. When considering all the minerals identified between 1990 and 2020, there is a systematic decline in the percentage that remain known only from their type locality that plots along an exponential curve:

$$\%\mathrm{M} = 10^{(-0.023\mathrm{Y}-0.01)}\ (r^2 = 0.9282)$$

where %M is the percentage of minerals from a year in the past Y that are known from only their type locality (Fig. 6a).

A second way to characterise the discovery of additional mineral locations is to calculate the average number of known localities for all minerals accepted by the IMA in a given year using the same snapshot of 1990 to 2020 data. In this case, a linear increase in the average number of known locations of minerals discovered in a given year with increasing age is observed (Fig. 6b). While dispersion around this trend increases with age, the relationship holds very strongly for minerals added to the database over the 15 years preceding the snapshot. The observed linear relationship is defined by the equation:

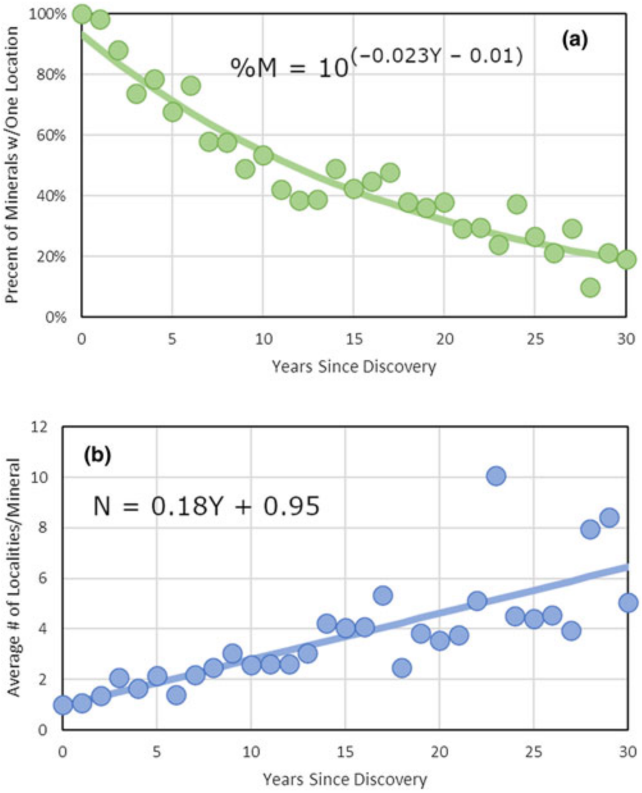

Figure 6. When first discovered, new mineral species are typically exclusively endemic to their type locality. Over time, most minerals are subsequently identified in other locations. (a) The percentage of minerals from a year in the past that are known only from their type locality decreases with increasing time describing an exponential curve [$\%\mathrm{M} = 10^{(-0.23\mathrm{Y}\ -\ 0.01)}$]. (b) Similarly, the average number of known locations for minerals increases with time since their discovery following a linear function [N = 0.18Y + 0.95] such that an additional location is found, on average, every 5.5 years (1/0.18). The *y*-intercept value of 0.95 confirms the notion that most minerals are initially known from only one location.

$$\mathrm{N}\ = 0.18\mathrm{Y} + 0.95\ \ (r^2 = 0.7871)$$

Where N is the average number of localities per mineral and Y is the years that have transpired since the acceptance of the mineral into the IMA database. Thus, on average, over the last thirty years, an additional locality of occurrence for recently discovered minerals is reported every 5.5 years (1/0.18 *y*), while the *y*-intercept value (0.95) supports the notion that newly discovered minerals are commonly known only from their type location. In both of these treatments, mineral abundance data has been normalised to the total number of minerals discovered in a given year in order to correct for temporal variation in the number of minerals identified annually.

When extending the analysis of minerals known from a single location back in time, both expected and unexpected results are observed. First, as expected, most endemic mineral species have

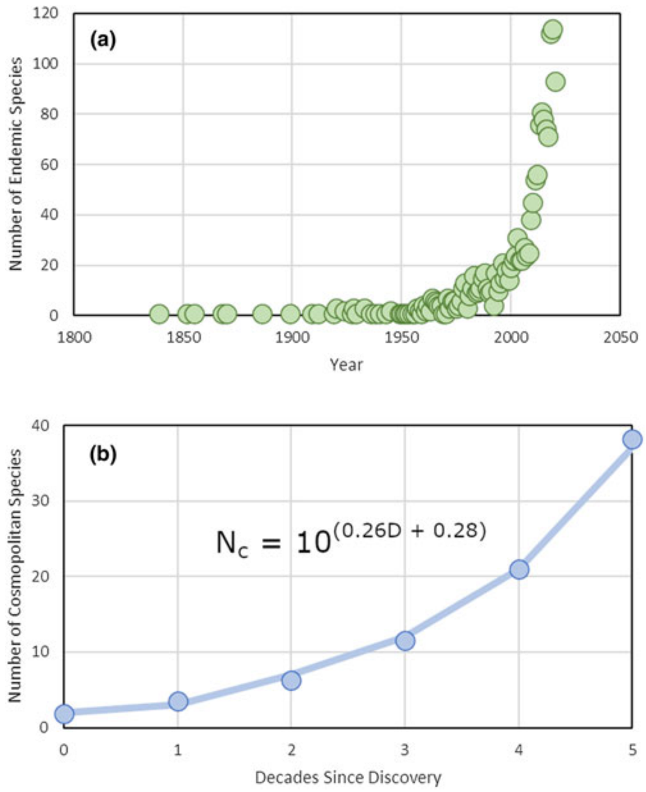

Figure 7. Most minerals that are known from only their type locality have been discovered recently. However, there are some species that were first discovered in the past that remain endemic to their type locality for extended periods (a). Conversely, some recently discovered minerals are now known from a relatively large number of locations. The number of cosmopolitan species (here defined as 50 or more mineral-location pairs) increases with time. When aggregating the number of such species per decade, this increase follows an exponential curve [$N_c = 10^{(0.26D + 0.28)}$] (b).

only recently been discovered (Fig. 7a). For the majority of these minerals, future workers can be expected to recognise their occurrence in other locations (Fig. 6a and b). Second, and perhaps quite unexpectedly, there is a group of minerals that were first recognised long ago that have continued to remain known from only a single location, the date of discovery of some of these species extends back to the 19th Century (Fig. 7a). Unlike the representation of endemic mineral species in Fig. 6a wherein the percentage of endemic minerals per year was used, this graph displays the absolute number of minerals that have remained endemic to their type locality since their discovery. Furthermore, since most widely distributed species have probably been discovered already, and as research focuses on known diversity hotspots with enhanced technology, the proportion of minerals with high endemicity is likely to increase.

These data suggest that it is unlikely that a previously unknown mineral will be discovered that comprises a significant volume of the Earth's lithosphere. However, workers will probably continue to find minerals that are subsequently identified at a relatively large number of locations globally. In order to evaluate this aspect of the distribution of mineral species those minerals with 50 or more known locations in the Mindat database are considered a cosmopolitan species. Though this is a somewhat arbitrary distinction based upon less than systematically obtained data, it is significant that the cutoff of 50 locations is a value which is significantly less than that point in the frequency of locations data where dispersion around the observed power-law linear relationship becomes pronounced (Fig. 4b). Because the discovery of both new minerals and new locations of previously known minerals is ongoing, for this analysis the January, 2021 snapshot of the Mindat database is used to tabulate the total number of cosmopolitan species that had first been identified in previous decades.

Of the 1,084 minerals recognised during the years 2011 to 2020, only two are found to have subsequently been recognised at 50 or more locations (clino-suenoite and 'eleonorite'). Over the preceding six decades the number of cosmopolitan mineral species identified per decade increases such that during the 1960s 37 minerals were identified that in 2021 fit the definition of a cosmopolitan species (Fig. 7b). As such, the number of cosmopolitan species per decade is found to follow the exponential growth function:

$$N_c = 10^{(0.26D + 0.28)} \ (r^2 = 0.9971)$$

Where N_c is the number of cosmopolitan species now recognised from previous decades D. An increase in the number of cosmopolitan species with age is perhaps not inherently surprising, however the highly systematic way in which the number grows suggests that a small but significant number of recently recognised species will ultimately be recognised as cosmopolitan through the efforts of future workers.

Conclusions

The heterogeneous distribution of global mineral abundances has been evaluated in several novel ways. First, the Earth's mineral hotspots exhibit an exponential distribution of species abundance while also exhibiting a power-law distribution in the number of minerals first recognised at those locations. Second, globally two-thirds of all known species have been reported from ten or fewer locations and the frequency distribution of these mineral-location pairs exhibit a power-law relationship that extends with increasing dispersion over several orders of magnitude of mineral occurrence abundance. Third, temporal trends in the identification of recently discovered minerals at additional locations have been shown to define relationships such that the percentage of recently discovered minerals that continue to remain known only from their type locality declines exponentially with time and the number of recently discovered minerals that are subsequently recognised at a large number of locations globally increases exponentially with time. On average recently discovered minerals are identified at new locations at a rate of about one new location every five and half years.

Acknowledgements. This work is an outgrowth from previous work on the hierarchical classification of fishes and other animal groups conducted with Bruce Wilkinson and Linda Ivany, both of whom made helpful suggestions that significantly enhanced this manuscript. I am also indebted to the students who enrolled in mineralogy at Purdue University Fort Wayne in the spring of 2022, their enthusiasm reinvigorated my interest in mineralogy. Finally, I would like to thank Dexter Perkins for patiently answering countless emails throughout this study. The work of reviewers and editors is gratefully acknowledged.

Competing interests. The authors declare none.

References

Alroy J., Marshall C.R., Bambach R.K., Bezuko K., Foote M., Fürsich F.T., Hansen T.A., Holland S.M., Ivany L.C., Jablonsky D., Jacobs D.K., Jones C., Kosnik M.A., Lidgard S., Low S., Miller A.I., Novack-Grottshall P.M., Olszewski T.D., Patzkowsky M.E., Raup D.M., Roy K., Sepkosik J.J., Somers M.G., Wagner P.J. and Webber A. (2001) Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Science*, **98**, 6261–6266.

Alroy J., Aberhan M., Bottjer D.J., Foote M., Fürsich F.T. Harries P.J., Hendy A.J.W., Holland S.M., Ivany L.C., Kiessling W., Kosnik M.A., Marshall C.R., McGowan A.J., Villier L., Wagner P.J., Bonuso N., Borkow P.S., Brenneis B., Clapham M.E., Fall L.M., Ferguson C.A., Hanson V.L., Krug A.Z., Layou K.M., Leckey E.H., Nürnberg S., Powers C.M., Sessa J.A., Simpson.C., Tomašovych A. and Visaggi C.C. (2008) Phanerozoic trends in the global diversity of marine invertebrates. *Science*, **321**, 97–100.

Atencio D. (2000) Minerals for which Brazil is the type locality. *Rocks & Minerals*, **75**, 44–46.

Back M.E. and Mandarino J.A. (1999) Minerals for which Mexico is the type locality. *Rocks & Minerals*, **74**: 12–14.

Baijot M., Hatert F. and Philippo S. (2012) Mineralogy and geochemistry of phosphates and silicates in the Sapucaia pegmatite, Minas Gerais, Brazil: Genetic implications. *The Canadian Mineralogist*, **50**, 1531–1554.

Bandy M.C. (1938) Mineralogy of three sulphate deposits of Northern Chile. *American Mineralogist*, **23**, 669–760.

Bermanec M., Vidovic N., Gavryliv L., Morrison S.M. and Hazen R.M. (2022) Evolution of symmetry index in minerals. *Geoscience Data Journal*, **2022**, 10.1002/gdj3.177.

Britvin S.N., Murashko M.N., Vereshchagin O.S., Vapnik Y., Shilovskikh V.V., Vlasenko N.S. and Permyakov V.V. (2022) Expanding the speciation of terrestrial molybdenum: Discovery of polekhovskyite, $MoNiP_2$, and insights into the sources of Mo-phosphides in the Dead Sea Transform area. *American Mineralogist*, **107**, 2201–2211.

Burke E.A.J. and Leake B.E. (2005) "Named amphiboles": a new category of amphiboles recognized by the International Mineralogical Association (IMA), and the proper order of prefixes to be used in amphibole names. *American Mineralogist*, **90**, 516–517.

Caraballo M.A., Michel F.M. and Hochella M.F. (2015) The rapid expansion of environmental mineralogy in unconventional ways: Beyond the accepted definition of a mineral, the latest technology, and using nature as our guide. *American Mineralogist*, **100**, 14–25.

Cassedanne J.P. and Baptista A. (1999) Famous Mineral Localities: The Sapucaia Pegmatite, Minas Gerais, Brazil. *Mineralogical Record*, **30**, 347–360.

Christy A.G. (2015) Causes of anomalous mineral diversity in the Periodic Table. *Mineralogical Magazine*, **79**, 33–49.

Cook R.B. (1978) Famous mineral localities: Chuquicamata, Chile. *The Mineralogical Record*, **9**, 321–333.

Corral A. and González Á. (2019) Power law size distribution in geoscience revisited. *Earth and Space Science*, **6**, 673–697.

Downs R.T. (2006) The RRUFF Project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals. *Program and Abstracts of the 19th General Meeting of the International Mineralogical Association, Kobe, Japan*, O03–13.

Drummond C.N. (1999) Bed-thickness structure of multi-sourced ramp turbidites: Devonian Brallier Formation, Central Appalachian Basin. *Journal of Sedimentary Research*, **69**, 115–121.

Dunn P.J. (1995) *Franklin and Sterling Hill New Jersey: The World's Most Magnificent Mineral Deposits*. The Franklin-Ogdensburg Mineralogical Society, Franklin, New Jersey, USA.

El Goresy A., Nagel K., Dominik B. and Ramdohr P. (1977) Fremdlinge: potential presolar material in Ca, Al rich inclusions of Allende. *Meteoritics*, **12**, 215–216.

Fedotov S.A. and Markhinin Y.K. (1983) *The Great Tolbachik Fissure Eruption: Geological and Geophysical Data, 1975–1976*. Cambridge University Press, Cambridge, UK, 341 pp.

Fuchs Louis H. (1971) Occurrence of wollastonite, rhönite, and andradite in the Allende meteorite. *American Mineralogist*, **56**, 2053–2068.

Gavryliv L., Ponomar V., Bermanec M. and Putiš M. (2022) The taxonomy of mineral occurrence rarity and endemicity. *The Canadian Mineralogist*, **60**, 731–758.

Grew E.S. and Hazen R.M. 2014. Beryllium mineral evolution. *American Mineralogist*, **99**, 999–1021.

Grew E.S., Hystad G., Hazen R.M., Krivovichev S.V. and Gorelova L.A. (2017) How many boron minerals occur in Earth's upper crust? *American Mineralogist*, **102**, 1573–1587.

Grey I.E. (2022) Diffraction methods in the characterization of new mineral species. *Journal of Solid State Chemistry*, **312**, 123239.

Hawthorne F.C., Oberti R., Harlow G.E., Maresch W.V., Martin R.F., Schumacher J.C. and Welch M.D. (2012) Nomenclature of the amphibole supergroup. *American Mineralogist*, **97**, 2031–2048.

Hazen R.M. and Ausubel J.H. (2016) On the nature and significance of rarity in mineralogy. *American Mineralogist*, **101**, 1245–1251.

Hazen R.M., Grew E.S., Downs R.T., Golden J. and Hystad G. (2015) Mineral ecology: Chance and necessity in the mineral diversity of terrestrial planets. *The Canadian Mineralogist*, **53**, 295–324.

Hazen R.M., Morrison S.M., Krivovichev S.V. and Downs R.T. (2022) Lumping and splitting: Towards a classification of mineral natural kinds. *American Mineralogist*, **107**, 1288–1301.

Hummer D.R. (2021) Fractal distribution of mineral species among the crystallographic point groups. *American Mineralogist*, **106**, 1574–1579.

Hystad G., Downs R.T. Hazen R.M. (2015a) Mineral species frequency distribution conforms to a large number of rare events model: Prediction of Earth's missing minerals. *Mathematical Geoscience*, **47**, 647–661.

Hystad G., Downs R.T., Grew E.S. and Hazen R.M. (2015b) Statistical analysis of mineral diversity and distribution: Earth's mineralogy is unique. *Earth and Planetary Science Letters*, **426**, 154–157.

Leake B.E. (1978) Nomenclature of amphiboles. *American Mineralogist*, **63**, 1023–1052.

Marchese C. (2015) Biodiversity hotspots: A shortcut for a more complicated concept. *Global Ecology and Conservation*, **3**, 297–309.

Markl G., Keim M.F. and Bayerl R. (2019) Unusual mineral diversity in a hydrothermal vein-type deposit: the Clara Mine, SW Germany, as a type example. *The Canadian Mineralogist*, **57**, 427–456.

Mindat (2023) *The Hudson Mineralogical Association*. https://www.mindat.org [accessed January 11, 2023].

Mücke A. (1981) The parageneses of the phosphate minerals of the Hagendorf pegmatite - a general view. *Chemie der Erde*, **40**, 217–234.

Myers N. (1988) Threatened biotas: "Hot spots" in tropical forests. *The Environmentalist*, **8**, 1–20.

Pasero M. (2023) *The New IMA List of Minerals*. International Mineralogical Association. Commission on new minerals, nomenclature and classification (IMA-CNMNC). http://cnmnc.units.it/ [IMA_Master_List_(2022–11).pdf, accessed January 11, 2023]

Pekov I.V., Agakhanov A.A., Zubkova N.V., Belakovskiy D.I., Vigasina M.F., Britvin S.N., Turchkova A.G. and Sidorov E.G. (2020) Philoxenite, $(K,Na,Pb)_4(Na,Ca)_2(Mg,Cu)_3(Al_{0.5})(SO_4)_8$, a new mineral from the fumarole exhalations of the Tolbachik Volcano, Kamchatka, Russia. *Zapiski Rossiiskogo Mineralogicheskogo Obshchetstva*, **149**, 67–77.

Pelloux A. (1927) The minerals of Vesuvius. *American Mineralogist*, **12**, 14–21.

Ponomar V., Gavryliv L. and Putiš M. (2023) The spatial and temporal evolution of mineral discoveries and their impact on mineral rarity. *American Mineralogist*, **108**, 1483–1494.

Reid W.J. (1998) Biodiversity hotspots. *Trends in Ecology and Evolution*, **13**, 275–280.

Rothman D.H., Grotzinger J.P. and Flemings P. (1994) Scaling in turbidite deposition. *Journal of Sedimentary Research*, **A64**, 59–67.

Russo M., Campostrini I. and Demartin F. (2014) Fumarolic minerals after the 1944 Vesuvius eruption. *Plinius*, **40**, 306.

Skinner B.J. and Skinner C.W. (1980) Is there a limit to the number of minerals? *The Mineralogical Record*, **11**, 333–335.

Turcotte D.L. (1994) Fractal aspects of geomorphic and stratigraphic processes. *GSA Today*, **4**, 201–213.

Addendum During production of this paper, it was brought to the attention of the editors that a number of other papers had been published recently which contained similar subject matter and data to those contained in the present work. The related publications were reviewed by experts and it was ascertained that they should have been cited by the present study. Citations of the papers by Gavryliv *et al.* (2022) and Ponomar *et al.* (2023) have now been added to the discussion and the related references added to the reference list. Interested readers are recommended to read these studies to explore further the topics investigated here.