Biota J. Wagner 2004 [J. Wiemann, K. de Queiroz, T. B. Rowe, N. J. Planavsky, R. P. Anderson, J. P. Gogarten, P. E. Turner, and J. A. Gauthier], converted clade name

Registration Number: 298

Definition: The largest crown clade containing *Homo sapiens* Linnaeus 1758. This is a special case of the maximum-crown-clade definition in that it does not use an external specifier (see Comments); it refers to the crown clade including humans and all other bioentities sharing common ancestry with them (see Comments). Abbreviated definition: max crown ∇ (*Homo sapiens* Linnaeus 1758).

Etymology: Latinized from Ancient Greek βίος (syncope of βίοτος) meaning "life".

Reference Phylogeny: Figure 1 in Hug et al. (2016), in which *Homo sapiens* is part of the clade named *Opisthokonta* (and *Biota* includes all taxa in that tree). See also Lake et al. (1984, "Eocytic Tree" in "Note Added in Proof" on p. 3790), Woese et al. (1990: Fig. 1), Pace (2006), Williams et al. (2012: Fig. 1a), Gouy et al. (2015: Fig. 3), and Hinchliff et al. (2015: Fig. 1).

Composition: Assuming a monophyletic origin of life on Earth (Theobald, 2010), this clade includes the last universal common ancestor (LUCA) of all "living" biological entities (Penny and Poole, 1999) on Earth, and all of its descendants, both extant and extinct, on this planet or anywhere else in the universe. Accumulating evidence for multiple derivations of the building blocks of *Biota* (see *Pan-Biota*, this volume) does not contradict a single origin of *Biota*. Moreover, even should multiple origins of life (on Earth or elsewhere) be demonstrated, if those life-forms do not share a genealogical connection to *Homo sapiens*, then they would not be members of *Biota* (see *Pan-Biota*, this volume).

At our current state of knowledge, *Biota* is generally thought to be composed of three major groups: *Bacteria*, *Archaea*, and *Eukarya* (= BAE), the first two of which might be paraphyla rather than clades (see Comments). In contrast, *Eukarya* (or its synonym *Eukaryota*) has always been associated with a major branch of the tree of life that is now universally regarded as a well-supported clade. Several new taxa of uncertain relations have recently been discovered; they are either deeply imbedded in *Bacteria* (Hug et al., 2016: Suppl. Fig. 2) or join the rest of *Biota* near the putative root of the biotan tree (i.e., some could be sister to all other biotans) (Hug et al., 2016: Suppl. Fig. 1). There are additional sets of (potential) biological entities, or parts of entities, that, while not themselves organisms as traditionally conceived (see Comments), are composed of potentially clade-forming entities that are often discussed when considering the transition from chemistry to biology (i.e., the origin of life), and might thus be parts of *Biota*: viruses, transposons, and nanobacteria (e.g., Trifonov and Kejnovsky, 2015; Kejnovsky and Trifonov, 2016; and see Comments). Extant viruses and transposons represent likely non-monophyletic groups of molecular parasites that seem to be parts of *Pan-Biota*, but not necessarily, or only in part, members of *Biota* (see Comments). Terrestrial and Martian nanobacteria have been suggested as members of *Biota*, posing a significant challenge to understanding the biogeography of early life. Because the phylogeny and composition of most

of these candidate biological entities (bioentities) are not well understood, we shall review current insights on each one of them and discuss their significance, composition, and phylogenetic placement in the Comments sections of the contributions on *Biota* and *Pan-Biota*.

If all extant viruses and transposons represent degenerate derivatives of the last common ancestor of bacteria, archaeans, and eukaryotes (BAE), then they are parts of *Biota* as currently understood. If some or all of those bioentities are derivatives of cellular pan-biotans outside of the BAE crown, or even earlier-diverging precellular assemblies that share common ancestry with BAE, they would still be parts of *Biota* as defined here, but that name would apply to a more inclusive clade than the BAE crown. Moreover, if some or all viruses and transposons are found to derive from the other branch stemming from the very earliest divergence of replicators leading to the crown, the name *Biota* would become synonymous with *Pan-Biota*. In each case, the Composition and Diagnosis of *Biota* would change accordingly.

Diagnostic Apomorphies: The definition of the taxon name *Biota* is clear in theory—it refers to the maximum crown clade containing humans, that is, to all descendants of the last common ancestor of all life on Earth today. In practical terms, however, the exact composition of this clade remains elusive, and its diagnosis accordingly unclear. Given that there is still much to be learned about the microorganisms of our planet, there is a possibility that the basal-most crown divergence in the tree of life on Earth has yet to be discovered (see, e.g., "pandoraviruses" sensu Philippe et al., 2013; but see contrary view by Koonin and Yutin, 2018). The diagnosis of *Biota* is further complicated by uncertainty regarding potential phylogenetic relationships of some classes of bioentities, such as viruses and transposons, which may be secondarily simplified molecular parasites derived from cellular (or even pre-cellular) lifeforms (see Comments).

We accordingly focus on the diagnosis of the clade stemming from the last common ancestor that *Homo sapiens* (*Eukarya*) shared with *Halobacterium salinarum* (*Archaea*) and *Escherichia coli* (*Bacteria*), fully cognizant that at least some and probably all of these apomorphies could have arisen well before that ancestor (see Comments below and *Pan-Biota*, this volume). If some or all viruses (or mobile genetic elements) turn out to be members of the BAE clade (see Composition), then they must have lost (most of) these diagnostic features.

The ancestral biotan cell reproduced by binary fission (see *Pan-Biota,* this volume). It was enclosed in an envelope composed of lipids and proteins, containing proteinaceous transmembrane units that facilitate chemosensing of environmental stimuli and endergonic osmoregulation, and chemiosmotic coupling (Oró et al., 1990; Gogarten and Taiz, 1992; Orgel, 1994; Maynard Smith and Szathmáry, 1995; Sára and Sleytr, 2000; Klug et al., 2017; Rodrigues-Oliveira et al., 2017). It stored all vital biomolecules, i.e., amino acid-based proteins (organismic function and structure), fatty acid-based lipids (organismic structure), ribonucleic acids (organismic function), deoxyribonucleic acids (organismic function), and most, if not all, of their monomeric building blocks in a water-based medium (Oró et al., 1990; Orgel, 1994). Unlike some molecular assemblies (e.g., some nanobacteria), the ancestral biotan was capable of growth and reproduction via internal biosynthesis including ribosomes translating genes into proteins, and higher-level compound modification (= metabolism), that responded to chemo-sensed environmental stimuli through intragenerational metabolic adjustments (= physiological adaptation) and intergenerational modifications (= evolution) (Maynard Smith

and Szathmáry, 1995, 1997). The ancestral biotan cell stored its DNA wrapped around proteins in chromosomes and contained a set of at least 355 genes (Martin et al., 2016).

Synonyms: Approximate (and partial if descendant viruses, mobile genetic elements, and nanobacteria are excluded) synonyms: *Zotica* Spinola 1850; *Bionta* Walton 1930; *Cellulata* Vorontsov 1965; *Cytobiota* Hu 1965; *Cytota* and *Acytota* Trifonov and Kejnovsky 2015. Numerous partial synonyms refer to a paraphyletic taxon originating in approximately the same ancestor, including *Monera* Haeckel 1866, *Protista* Haeckel 1866 (but not as used by Copeland, 1938, and others), *Bacteria* Haeckel 1894, *Prokaryota* Swain 1969, and *Procytota* Jeffrey 1971 (see https://species.wikimedia.org/wiki/Prokaryota for a more complete list).

Comments: Given the fundamental significance of the most inclusive crown clade of life on Earth, it is surprising that the literature records so few instances in which it has been referred to explicitly by name, rather than being described in terms of the properties thought necessary to its existence—i.e., the often vague concept of "being alive" (see *Pan-Biota*, this volume)—or referred to by implication, for example, through reference to LUCA, the last universal common ancestor (Woese, 1999). By contrast, an inordinate amount of literature has been devoted to how many primary subtaxa should be recognized within this (seldom-named) taxon, and which traditional categorical ranks (e.g., Domain, Kingdom, Phylum) should be assigned to them (e.g., Blackwell, 2004, and references therein). That said, there are a number of taxon names that could arguably be applied to the crown clade that contains all bioentities originating on Earth (related to humans). Most of these candidate names were proposed explicitly for the paraphylum composed of cells lacking a membrane-bound nucleus (first distinguished by Chatton, 1925; see also Chatton, 1937/1938; though not christened formally with the name "*Prokaryota*" until Swain, 1969). Also, there are a number of candidate names associated with paraphyla that explicitly exclude multicellular organisms (e.g., *Protista* Haeckel 1866; in the twentieth century, this name became associated with unicellular eukaryotic organisms).

There is a much shorter list of taxon names, including *Zotica* Spinola 1850, *Bionta* Walton 1930, *Cytobiota* Hu 1965, *Cellulata* Vorontsov 1965, and *Cytota* Trifonov and Kejnovsky 2015 proposed for a group of bioentities including all cells, whether or not they have a membrane-bound nucleus, as well as both unicellular and multicellular organisms (although still potentially paraphyletic as this group often excluded viruses). Most biologists are unlikely to be familiar with most of these names for they have seldom been used since they were coined.

Biota is an exception, albeit best known as a common noun rather than as a taxon name. Stejneger (1901: 89) first proposed "biota" "as a comprehensive term to include both fauna and flora that will not only designate all animal and plant life of a given region or period, but also any treatise upon the animals and plants of any geographical area or geological period." Following Stejneger, its general use has been as a common noun and "biota" is usually qualified with an adjective (e.g., marine biota, extinct biota), with a prefix (e.g., microbiota, symbiota), or in combination with a possessive (e.g., biota of Brazil). Originally focused on multicellular organisms (i.e., terrestrial flora and fauna), it has since been expanded to include microbes in diverse settings (e.g., marine microbiota or gut biota). Nevertheless, when used as a common noun, "biota" generally refers to some subset of the maximum-crown-clade that is usually not itself a clade.

The term appears only recently to have been employed as a proper noun referring to the taxon composed of all bioentities on Earth. *Biota* was apparently first used in its taxonomic sense in a website (Brands, 1989–2005), then in an abstract volume produced for the Paris meeting of the International Society of Phylogenetic Nomenclature (Wagner, 2004), and later still in a self-published book (Pelletier, 2015). The earliest peer-reviewed journal article using *Biota* as a formal taxon name appears to be that of Trifonov and Kejnovsky (2015; see also Kejnovsky and Trifonov, 2016) (although Hu [1965] also used the term "biota" for "an immense group of living beings" [p. 255; including all viruses regardless of their likely polyphyly]). Given that Wagner (2004) was the first to apply *Biota* as a formal taxon name to this clade in a printed work (which can be downloaded and printed at the phylocode.org website), thereby qualifying it as "in use" and therefore a preexisting name under Article 6.2(b) of the *PhyloCode*, Wagner (2004) is here regarded as the nominal author of *Biota*.

Selecting *Biota* from among the seldom-used names available for this clade takes advantage of the implicit connection between this familiar term for "living beings" and the most inclusive crown clade originating on Earth. Whether used in its ecological or taxonomic sense, expressions such as the "biota" of South America or "biota" of the Burgess Shale would remain perfectly intelligible as referring to all organisms from those times and places. Moreover, as Stejneger noted, "*Biota*" has an obvious (etymological) connection to "biology", the name for the science of life. We accordingly take this opportunity to propose an explicit phylogenetic definition for the taxon name *Biota* using a modification of the form of definition proposed by Wagner (2004), i.e., a maximum-crown-clade definition that does not employ an external specifier. This departure from the standard form of a maximum-crown-clade definition reflects the exceptional circumstance that the name *Biota* is being applied to *the most* inclusive crown clade on this planet, rendering an external specifier superfluous to limiting the inclusiveness (circumscription) of the clade.

The single internal specifier, *Homo sapiens*, possesses all of the qualities most desirable in (if not required of) a nomenclatural specifier: it is composed of abundant, large organisms that are easily dissected, thoroughly studied and exhaustively illustrated from histological to gross anatomical levels, well-known ontogenetically, physiologically, and behaviorally, and it has a completely sequenced genome and an informative fossil record. It is also deeply imbedded within the named clade, thus buffering the name against unintended taxonomic consequences caused by potential changes in tree topology (e.g., Sereno, 1998). All extant organisms are no less distant in time than *Homo sapiens* is from the most recent common ancestor (= ur-ancestor) of *Biota*, were no less subject to the cumulative contingencies of evolution than we were, and are no less interesting and informative in their own rights (Gee, 2013). Nonetheless, biology is fundamentally and uniquely a human enterprise, as is the practice of coining taxon names to enhance cognitive efficiency (Mervis and Rosch, 1981) as humans communicate about *Biota* and its members. Thus, our specifier choice is not anthropocentric in a prejudicial or pejorative sense; rather, it employs the one and only species in all of *Biota* that is universally known to those creatures who study it, while underscoring that humanity's place in nature derives from a phylogenetic connection with all other life on Earth.

By drawing a clear distinction between the largest crown clade containing *Homo sapiens* (*Biota*) and its corresponding total clade (*Pan-Biota*, this volume), we hope to sharpen the focus of questions regarding the origin of the

last common ancestor of all extant replicators related to *H. sapiens*, the relationships of disparate "non-cellular" bioentities (e.g., viruses), and the origin of biological replication itself (sensu *Pan-Biota*).

Tying the taxon name *Biota* to the maximum crown clade containing *Homo sapiens* ensures that all extant organisms and other (less complex) extant bioentities sharing common ancestry with them will be parts of that clade. Nevertheless, ideas about the circumscription of the clade *Biota* may expand stem-ward in the future as new bioentities are discovered, and if some or all debated members of *Biota*, such as viruses and mobile genetic elements, are inferred to have emerged from bioentities that diverged before the last ancestor shared by *E. coli*, *H. salinarum*, and *H. sapiens*, and persisted to the Recent by interacting symbiotically with the descendants of that ancestor (see below). The basalmost possible phylogenetic position of the ancestor in which *Biota* originated would make this clade-name synonymous with *Pan-Biota* (this volume).

In the context of our definition for the name *Biota* and its proposed diagnosis, we now consider more controversial members of *Biota*, such as viruses, transposons, and nanobacteria, which are frequently discussed in modern literature addressing the origin of life. The *PhyloCode* governs clade nomenclature, and clades are typically thought to be composed of integrated bioentities (e.g., organisms) that are self-organizing, self-replicating, and self-sustaining; we accept that none of the controversial candidate members of *Biota* are "organisms" in that sense (Moreira and López-García, 2009). Nevertheless, we see no compelling reason to exclude them *a priori* from *Biota* as a clade (i.e., a common ancestor and *all* of its descendants), especially as they could represent simplified symbiotic offshoots of cellular (or pre-cellular) organisms (e.g., viruses) or their genomes (e.g., transposons and other mobile genetic elements). The following is a summary of key features of each of these potential members of *Biota*, a critical discussion of their significance to the origin of life, and an overview of current hypotheses concerning their phylogenetic placement, and how their inclusion in *Biota* might affect ideas about the composition of this clade now or in the future.

"Virus" is a functional term for a disparate array of DNA- or RNA-based replicators that are, with few exceptions (see below), enclosed in a proteinaceous sheath called a capsid; the term does not refer to a taxon as such (Van Regenmortel, 2003). These cell-reliant bioentities are the most abundant denizens of Earth, with at least 10 individual virus particles for every cell (Brüssow, 2009) (note also that your body contains about as many bacterial as human cells [Sender et al., 2016]), and have been found everywhere their hosts can survive, even in the most hostile environments (e.g., Sahara Desert; Prestel et al., 2012). Viral genomes typically contain two functional modules: those genes governing genome replication and those regulating capsid self-assembly (Krupovic and Koonin, 2017a,b). Indeed, bacteriophage genomes and capsids are often assembled separately in the cytoplasm of their hosts, where they must encounter one another in order for the former to insert into the latter. In that sense, viruses appear to be composite bioentities, analogous to eukaryotes (or, for that matter, all cells). As Krupovic and Koonin (2017a) argued, the capsid is the key innovation distinguishing viruses from among other "selfish" genetic elements, protecting their replicators in extracellular environments and enabling them to attach to and enter host cells.

Viruses are often explicitly excluded from taxa composed of typical (cellular) organisms bearing the diagnostic features noted above (see Diagnosis), including *Bionta* Walton 1930, *Cellulata* Vorontsov 1965, and *Cytobiota*

Hu 1965 (viruses had yet to be discovered when *Zotica* Spinola 1850 was proposed). In contrast, Trifonov and Kejnovsky (2015; see also Hu, 1965) explicitly included viruses in their *Biota* (as "living beings" if not as "organisms", the latter of which were included in their purportedly less-inclusive *Cytota* Trifonov and Kejnovsky 2015; see also Kejnovsky and Trifonov, 2016). It is not clear whether some or all viruses represent incipient or vestigial organisms, or whether they are rogue bits of genomes that escaped from cellular life-forms (see, e.g., Luria and Darnell, 1967; Podolsky, 1996; Forterre, 2006; Nasir and Caetano-Anollés, 2015), or some combination of these scenarios (e.g., Krupovic and Koonin, 2017a,b).

Viruses are readily distinguished from all other bioentities; for example, their capsids are morphologically and compositionally distinct from cell membranes, and their genome-replicating proteins lack any clear homologs among cellular life-forms (Krupovic and Koonin, 2017b). They are unable to replicate without their hosts, and for that reason it has often been questioned whether viruses are "alive" (see Moreira and López-García, 2009 and López-García and Moreira, 2009, and references therein). Such reasoning seems flawed, however, as many cellular parasites are unable to replicate without their hosts (e.g., *Mycobacterium tuberculosis*), yet they are nonetheless commonly considered to be "alive" (Hegde et al., 2009; Nasir and Caetano-Anollés, 2015). Granted that viruses differ from cells in that they are "unable to transform energy and matter (that is, to actively generate order from disorder)" (López-García and Moreira, 2009:15). But cellular genomes are likewise unable to accomplish those activities on their own and do not differ from viral genomes in that respect; both genomes require access to the elaborate metabolic machinery—including other genes, transcription products, proteins, and metabolites—found only inside membrane-bound cells, in order to reproduce themselves. Although questions about the origin of life remain subject to lively debate (e.g., Root-Bernstein and Root-Bernstein, 2015), our more modest goal is to consider whether one or more viral clades could be descended from uncontroversial "living" bioentities, and thus be part(s) of *Biota*. All viruses sharing ancestry with *Homo sapiens* are by definition part of *Pan-Biota* because they fulfill the requirement of "biological replication" (if that evolved only once), viz., the production of descendants that are themselves able to reproduce (see *Pan-Biota*, this volume).

Inferring phylogenetic relationships among viruses and assessing their place(s) in the tree of life have proven challenging. The viral lifestyle could have originated more than 3.5 billion years ago (see below) and viral "generation times" (from entry to exit from host) can be measured in minutes. As a consequence, viruses can achieve extraordinary rates of evolutionary change that, combined with high mutation rates, can result in remarkably divergent gene sequences, problems for sequence alignment, and a propensity for long-branch attraction, all of which can confound phylogenetic inference. Patterns of viral descent can be obfuscated by additional phenomena, such as horizontal gene transfer from host to virus, and further complicated by horizontal virus transfer among distantly related hosts (e.g., Dolja and Koonin, 2018). Despite these difficulties, several well-supported viral clades have been recognized in recent years, such as bacteriophage DNA viruses (Krupovic and Koonin, 2017a; Yang et al., 2018), eukaryote RNA viruses (Krupovic and Koonin, 2017b; Wolf et al., 2018), and the therian mammal *Lentivirus* clade (Nakano et al., 2017; which includes the subclade SIV and its subclade HIV [Sharp and Hahn, 2011]). No evidence has yet emerged indicating that all viruses comprise a single branch of the tree of

life (indeed, no single sequenceable biomolecule is shared by all viruses). Moreover, some viral capsid proteins appear to derive from proteins taken from different major subclades of cellular life-forms. This suggests that viruses—mobile genetic elements that can travel from cell to cell encased in protective sheaths derived from proteins transferred horizontally from preexisting cells—emerged on at least 20 separate occasions in the history of the biotan total clade (Krupovic and Koonin, 2017a; see also Gladyshev and Arkhipova, 2011).

It would be a mistake to regard any of these viruses as not being referable to *Biota* because they are thought, in some sense, not to be "alive". Crown caecilians (vermiform amphibians), for example, are still part of *Tetrapoda* even though they lack any vestiges of this clade's diagnostic four limbs (even as embryos). Similarly, even if viruses lack one, many, or all of the properties described in our Diagnosis, according to our definition, any given Recent virus stemming ultimately from a bioentity that shares ancestry with human cells would still be part of *Biota*, even if it diverged before the origin of the ancestral cell, or before the last cell shared by crown BAE (e.g., Nasir and Caetano-Anollés, 2015). Conversely, if a viral replicator originated independently (e.g., in an "RNA world" sensu Gilbert, 1986; Diener, 1989; Wolf et al., 2018), so that its genome did not share ancestry with *Homo sapiens* (e.g., Koonin and Dolja, 2014), then that virus would not be part of *Biota* (or *Pan-Biota*). That being said, if that independent viral line shared genes acquired from cellular life-forms via horizontal transfer (e.g., Krupovic and Koonin, 2017a), then there are some senses in which it could be considered part of *Biota* (*PhyloCode* Note 2.1.3). If an ancient pan-biotan virus parasitized some non-BAE host, it would only be part of *Biota* if that host was a direct ancestor of BAE, or if the virus later transferred horizontally to a BAE cell (and the virus survived to the Recent). On that note, some so-called "naked viruses" may well have evolved before the emergence of cellular life, and persisted in solution until they launched an evolutionary arms race with cellular bioentities (that likely predate the complex cell shared by the last common ancestor of BAE). If, as suggested by Koonin and Dolja (2014), viruses emerged from selfish genetic elements (and *vice versa*) on multiple occasions during evolution, then their inclusion in *Biota* would depend on the genealogical connections of those selfish genetic elements: do they ultimately share ancestry with uncontroversial biotan DNA-based genomes or not? Pinpointing the exact phylogenetic placement of the *Biota* node requires a clearer understanding of the origins and evolution of viral bioentities.

Transposons (short for "transposable elements"; also known as "jumping genes") do not comprise a taxon, but instead represent a functional class of DNA- or RNA-based replicators that can change position within the genome (transposition) through excision and insertion into a target DNA sequence (McClintock, 1950). These clade-forming bioentities, like other mobile genetic elements (MGEs, including self-splicing inteins and introns) spread—by both vertical and horizontal transfer—across the tree of life and can be grouped by various criteria, such as the distinctive catalytic mechanisms they employ during transposition (e.g., McDowall, 2006). Their propensity to proliferate within genomes—e.g., they comprise 42% of the human genome (Lander et al., 2001) and 85% of the maize genome (Schnable et al., 2009)—often with deleterious consequences, led to transposons being viewed as "selfish" genome parasites (in contrast to cell-parasitizing viruses). Their abundance within genomes doubtless creates more opportunities for horizontal gene transfer, especially by superabundant retrotransposons in eukaryotes (Kidwell, 1992), that not only jump about within their

genomes, but also generate numerous copies of themselves within the genome. Nevertheless, there is growing evidence that transposons may play important roles in the evolution of gene regulation in biotans (e.g., Wagner, 2014; Lanciano and Mirouze, 2018), offering new insights into how the functional morphology of the genome itself evolved (e.g., Britten and Davidson, 1971; Biémont, 2010).

Like viruses, mobile genetic elements are entirely dependent upon their hosts' metabolism and cannot replicate except inside cells (Kidwell, 1992; Villarreal, 2005). Questions about the evolutionary origin(s) of one class of MGEs, transposons, arose when some were found to share aspects of genome form and function with viruses (e.g., Villarreal, 2005). In particular, retroviral gene sequences have been identified in MGEs in eukaryotes that lack capsids and for that reason they have been regarded by some as "retrotransposons" (Koonin and Dolja, 2014). This suggested a possible evolutionary connection between these reverse-transcribing MGEs, with retrotransposons acquiring protein-coding genes from eukaryotic hosts to construct the viral capsids enabling these mobile genetic elements to move not just within cells, but also between them. This transition could have worked both ways, with retrotransposons potentially giving rise to retroviruses and *vice versa* on multiple occasions. However, Krupovic and Koonin (2017b) inferred that the capsidless "retrotransposons" currently known in *Eukarya*, such as non-mobile genetic elements residing exclusively in mitochondria, are actually secondarily simplified retroviruses, as they appear deeply imbedded within a very large eukaryote-wide clade of RNA-based viruses with fully developed viral capsids. An alternative hypothesis by Hickson (1989) suggests that some transposable genetic elements may have emerged from group-II-introns, short genetic elements capable of self-excision from their DNA strands.

Ideas about the evolutionary origins of transposons became intertwined with those of viruses. Nevertheless, whether transposons, or any other mobile genetic element, arose on one or more occasions before, during, or after the ancestral BAE cell first appeared, no one has yet presented any evidence that they are composed of anything but standard genes, such as those encoding transcriptase, common to all biotan cells (however modified these mobile elements might be compared to typical "immobile" genes). Thus, all transposons found in uncontroversial biotan genomes, regardless of their evolutionary origins, must be regarded as parts of *Biota* (and *Pan-Biota*).

Nanobacteria pose another challenge. As with "virus", the term does not apply to a taxon; it refers instead to tiny, spherical assemblies of biomolecules that attract mineral precipitation onto their surfaces. Comparable structures are widely distributed, having been extracted from the lithosphere, human kidney stones and atherosclerotic plaques, as well as from a Martian meteorite (as fossil remains). Whether nanobacteria (originally described as "ultramicrobacteria" by Torrella and Morita, 1981) can be considered "alive", and therefore potentially part of *Biota*, is controversial. Terrestrial and putative Martian nanobacteria (McKay et al., 1996; Çiftçioğlu et al., 2005; Martel and Young, 2008) share a mineral-encapsulated, multilayered organic sphere of very small size (~200 nm circumference), and are allegedly able to replicate this morphology (Sommer et al., 2003; Çiftçioğlu et al., 2005; Martel and Young, 2008) (although this could result from initiation of mineral precipitation, as binary fission has yet to be observed). The idea that nanobacteria are "alive" has been challenged by a fundamental reinterpretation of these structures as organic, nanoparticulate agglomerates that facilitate secondary mineral

precipitation via Coulomb effects, or coordination chemistry with surrounding dissolved ions (Cisar et al., 2000; Martel and Young., 2008). Nanobacteria extracted from biotans stain positively for DNA (which might be the product of probe/instrument contamination) (Cisar et al., 2000). However, putative Martian nanobacteria contain only polycyclic aromatic hydrocarbons that could still be fossilization products of original biomolecules (McKay et al., 1996; Wiemann et al., 2018). Therefore, nanobacteria would only be part of *Biota* if they descended from a pre-cellular ancestor of BAE cells, or are derived from secondarily simplified BAE cells. Or they might not be taxa *per se*, but merely represent biotan cellular debris acting as nuclei for mineral deposition, as might be the case for examples extracted from biotan host tissues.

While nanobacteria within human hosts could be part of crown *Biota* (e.g., Çiftçioğlu et al., 1999), putative Martian nanobacteria pose a special problem: their fossil nature suggests that these potentially biological entities existed multiple millions of years before present, well before any possibility of anthropogenic contamination of Earth's adjacent satellite bodies and planets. In that case, Martian nanobacteria might represent either non-anthropogenic contaminants from Earth, transported through whatever means (e.g., by impact debris ejected into space from asteroid collisions). Or they could represent a separate origin of extraterrestrial life, in which case they would not be part of *Biota* as defined here (see *Pan-Biota*, this volume). Alternatively, they could share a common ancestral replicator with *Homo sapiens*—transported either from Earth, from Mars, or to both planets from elsewhere—and thus be members of *Biota* or non-biotan *Pan-Biota*. Or they could just be non-living organic agglomerates arising through abiotic chemistry, and having nothing to do with *Pan-Biota* (although they could represent non-replicating mineralized biological detritus produced by members of that clade).

One reason for the numerous controversies around these potential candidate members of *Biota* is the fact that the biotan tree is difficult to root (e.g., Hug et al., 2016). It is generally thought that first *Archaea* (whether a clade or paraphylum), and then *Bacteria* (whether a clade or paraphylum), are successive sisters to *Eukarya* (the monophyly of which has never seriously been questioned since Chatton [1925] recognized their foremost shared apomorphies: a membrane-bound nucleus and mitochondria)—in other words, the root of *Biota* is widely held to lie either within (paraphyletic) "*Bacteria*" or between (monophyletic) *Bacteria* on the one hand, and an "*Archaea*" + *Eukarya* clade on the other (e.g., Gogarten et al., 1989; Iwabe et al., 1989; Ciccarelli et al., 2006; Fournier and Gogarten, 2010). As was the case for discovering the clade *Eukarya*, the idea that some "archaeans" are closer to eukaryotes than are others (the "Eocyte Hypothesis") was based initially on (micro)morphology (Lake et al., 1984), and both hypotheses have been corroborated in several recent gene-sequence-based studies (e.g., Williams et al., 2013; Gouy et al., 2015; Hug et al., 2016; Betts et al., 2018). Unfortunately, none of the potential candidate members of *Biota* considered here were included in any of these phylogenetic analyses due to biomolecular incompatibilities, or the current lack of phylogenetically informative molecular residues in their fossil representatives.

Based on our current understanding of *Biota*, it is thought that the ancestral biotan was heterotrophic (Blankenship, 2010; see also Betts et al., 2018) and must have evolved in an aqueous solution containing essential sodium, potassium, magnesium, chloride, phosphate, carbonate, sulfate, nitrate, transition metal ions, and the dissolved organic molecules upon which it

fed (e.g., Hunding et al., 2006). This ancestral habitat would not necessarily have to be located in the photic zone, as the primary environmental cue governing metabolic rhythmicity ancestral for *Biota* (e.g., convection—based on temperature change in the upper water layers, and/or UV-light-based chemical reactivity and redox gradients) could be sensed even in the aphotic zone. Biotans can be found today all over Earth's surface and shallow subsurface, and in its hydrosphere and atmosphere. Multiple species can now also be found in some of Earth's adjacent satellite bodies through space-travel-mediated transport of anthropogenic contaminants (Novikova et al., 2006).

Microbial fossils from the 1.9 billion-year-old Belcher Supergroup in Canada (Hofmann, 1976; Golubic and Hofmann, 1976) can unambiguously be identified as (*Pan-*) *Cyanobacteria*, and this clade must be older than the first permanent rise of atmospheric oxygen for which it was responsible—referred to as the Great Oxidation Event—at 2.45–2.32 billion years ago (e.g., Luo et al., 2016). Thus, the basal biotan divergence took place no less than 2.32–2.45 billion years ago based on a literal reading of the body-fossil record. However, the evolution of biotic oxygen production is typically inferred, based on geochemical proxies for divalent oxygen, to have occurred prior to 3.0 billion years ago (e.g., Wang et al., 2018). If isotopic or biomarker signatures correlated with methanogenesis were produced either by methanogenic archaeans, or by the ancestor of archaeans and eukaryotes (e.g., Betts et al., 2018), then *Biota* would be at least 3.5 billion years old. There are multiple claims based on biomarkers and isotope signatures suggesting that *Biota,* or at least *Pan-Biota,* predates 3.5 billion years, but these claims are debated. Divergence-time estimates based on molecular-clock analyses using ancestral biotan genes—those coding for proteins involved in binary fission, ATP metabolism, protein synthesis (cell division; housekeeping; metabolism), and protein digestion (heterotrophic lifestyle; primitive immune response)—suggest that the basal biotan divergence took place much earlier, more than 3.9 billion years ago (Betts et al., 2018). However, inferred divergence times for organisms passing through mass extinctions—and *Biota* has survived several—can distort such estimates (e.g., Berv and Field, 2018). Moreover, this estimate of divergence time would place the biotan origin before the end of the Hadean Heavy Bombardment, an interval in Earth history unlikely to have been favorable to the proliferation of delicate cellular life-forms. Nevertheless, modeling indicates that this catastrophic episode in Earth history was unlikely to have produced conditions exceeding biotic tolerances planet-wide (Abramov and Mojzsis, 2009). The origin of crown *Biota* is thus likely to have occurred sometime between these two estimates—between 3.9 Ba and 3.5 Ba—and the origin of the biotan total clade is likely to be much older (see *Pan-Biota*, this volume), though doubtless after the moon-forming impact 4.51 billion years ago thought to have sterilized the Earth's surface.

Literature Cited

Abramov, O., and S. J. Mojzsis. 2009. Microbial habitability of the Hadean Earth during the late heavy bombardment. *Nature* 459:419–422.

Berv, J. S., and D. J. Field. 2018. Genomic signature of an avian Lilliput Effect across the K-Pg extinction. *Syst. Biol.* 67(1):1–13.

Betts, H. C., M. N. Puttick, J. W. Clark, T. A. Williams, P. C. Donoghue, and D. Pisani. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* 2(10):1556–1562.

Biémont, C. 2010. A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* 186:1085–1093.

Blackwell, W. H. 2004. Is it kingdoms or domains? Confusion & solutions. *Am. Biol. Teach.* 66(4):268–276.

Blankenship, R. E. 2010. Early evolution of photosynthesis. *Plant Physiol.* 154(2):434–438.

Brands, S. J. 1989–2005. *Systema Naturae 2000*. Amsterdam, the Netherlands. Available at http://sn2000.taxonomy.nl/.

Britten, R. J., and E. H. Davidson. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q. Rev. Biol.* 46(2):111–138.

Brüssow, H. 2009. The not so universal tree of life *or* the place of viruses in the living world. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364:2263–2274.

Chatton, E. 1925. *Pansporella perplex*: reflections on the biology and phylogeny of the *Protozoa*. *Ann. Sci. Nat. Zool.* 8:5–85.

Chatton, E. 1937/1938. *Titres et Travaux Scientifiques (1906–1937)*. Sottano Imprimerie, Sète, France.

Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel, and P. Bork. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311(5765):1283–1287.

Çiftçioğlu, N., M. Björklund, K. Kuorikoski, K. Bergström, and E. O. Kajander. 1999. Nanobacteria: an infectious cause for kidney stone formation. *Kidney Int.* 56(5):1893–1898.

Çiftçioğlu, N., R. S. Haddad, D. C. Golden, D. R. Morrison, and D. S. McKay. 2005. A potential cause for kidney stone formation during space flights: enhanced growth of nanobacteria in microgravity. *Kidney Int.* 67(2):483–491.

Cisar, J. O., D. Q. Xu, J. Thompson, W. Swaim, L. Hu, and D. J. Kopecko. 2000. An alternative interpretation of nanobacteria-induced biomineralization. *Proc. Natl. Acad. Sci. USA* 97(21):11511–11515.

Copeland, H. F. 1938. The kingdoms of organisms. *Q. Rev. Biol.* 13:383–420.

Diener, T. O. 1989. Circular RNAs: relics of precellular evolution? *Proc. Natl. Acad. Sci. USA* 86(23):9370–9374.

Dolja, V. V., and E. V. Koonin. 2018. Metagenomics reshapes the concepts of RNA virus evolution by extensive horizontal virus transfer. *Virus Res.* 244:36–52.

Forterre, P. 2006. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* 117(1):5–16.

Fournier, G. P., and J. P. Gogarten. 2010. Rooting the ribosomal tree of life. *Mol. Biol. Evol.* 27(8):1792–1801.

Gee, H. 2013. *The Accidental Species: Misunderstandings of Human Evolution*. University of Chicago Press, Chicago, IL.

Gilbert, W. 1986. The RNA world. *Nature* 319:618.

Gladyshev, E. V., and I. R. Arkhipova. 2011. A widespread class of reverse transcriptase-related cellular genes. *Proc. Natl. Acad. Sci. USA* 108(51):20311–20316.

Gogarten, J. P., H. Kibak, P. Dittrich, L. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, T. Date, and T. Oshima. 1989. Evolution of the vacuolar H+-ATPase: implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* 86(17):6661–6665.

Gogarten, J. P., and L. Taiz. 1992. Evolution of proton pumping ATPases: rooting the tree of life. *Photosynth. Res.* 33:137–146.

Golubic, S., and H. J. Hofmann. 1976. Comparison of modern and mid-Precambrian *Entophysalidacae* (*Cyanophyta*) in stromatolitic algal mats: cell division and degradation. *J. Paleontol.* 50:1074–1082.

Gouy, R., D. Baurain, and H. Philippe. 2015. Rooting the tree of life: the phylogenetic jury is still out. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140329.

Haeckel, E. 1866. *Generelle Morphologie der Organismen*. Georg Reimer, Berlin.

Haeckel, E. 1894. *Systematische Phylogenie der Protisten und Pflanzen*. Georg Reimer, Berlin.

Hegde, N. R, M. S. Maddur, S. V. Kaveri, and J. Bayry. 2009. Reasons to include viruses in the tree of life. *Nat. Rev. Microbiol.* 7(8):615.

Hickson, R. E. 1989. Self-splicing introns as a source for transposable genetic elements. *J. Theor. Biol.* 141(1):1–10.

Hinchliff, C. E., S. A. Smith, J. F. Allman, J. G. Burleigh, R. Chaudhary, L. M. Coghill, K. A. Crandall, J. Deng, B. T. Drew, R. Gazis, and K. Gude. 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl. Acad. Sci. USA* 112(41):12764–12769.

Hofmann, H. J. 1976. Precambrian microflora, Belcher Islands, Canada—significance and systematics. *J. Paleontol.* 50:1040–1073.

Hu, H. H. 1965. The major groups of living beings: a new classification. *Taxon* 14:254–261.

Hug, L. A., B. B. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Y. Amano, K. Ise, Y. Susuki, N. Dudek, A. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, and J. F. Banfield. 2016. A new view of the tree of life. *Nat. Microbiol.* 1:16048.

Hunding, A., F. Kepes, D. Lancet, A. Minsky, V. Norris, D. Raine, K. Sriram, and R. Root-Bernstein. 2006. Compositional complementarity and prebiotic ecology in the origin of life. *Bioessays* 28(4):399–412.

Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata. 1989. Evolutionary relationship of *Archaebacteria*, *Eubacteria*, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA* 86(23):9355–9359.

Jeffrey, C. 1971. Thallophytes and kingdoms: a critique. *Kew Bull.* 25(2):291–299.

Kejnovsky, E., and E. N. Trifonov. 2016. Horizontal transfer-imperative mission of acellular life forms, *Acytota*. *Mob. Genet. Elements* 6(2):1154636.

Kidwell, M. G. 1992. Horizontal transfer of P elements and other short inverted repeat transposons. *Genetica* 86(1–3):275–286.

Klug, Y. A., E. Rotem, R. Schwarzer, and Y. Shai. 2017. Mapping out the intricate relationship of the HIV envelope protein and the membrane environment. *Biochim. Biophys. Acta (BBA)—Biomembr.* 1859(4):550–560.

Koonin, E. V., and V. V. Dolja. 2014. Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol. Mol. Biol. Rev.* 78(2):278–303.

Koonin, E. V., and N. Yutin. 2018. Multiple evolutionary origins of giant viruses. *F1000Research* 7:1840.

Krupovic, M., and E. V. Kroonin. 2017a. Multiple origins of viral capsid proteins from cellular ancestors. *Proc. Natl. Acad. Sci. USA* 14(12):2401–2410.

Krupovic, M., and E. V. Kroonin. 2017b. Homologous capsid proteins testify to the common ancestry of *Retroviruses*, *Caulimoviruses*, *Pseudoviruses*, and *Metaviruses*. *J. Virol.* 9(12):e00210–17.

Lake, J. A., E. Henderson, M. Oakes, and M. W. Clark. 1984. Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. *Proc. Natl. Acad. Sci. USA* 81(12):3786–3790.

Lanciano, S., and M. Mirouze. 2018. Transposable elements: all mobile, all different, some stress responsive, some adaptive. *Curr. Opin. Genet. Dev.* 49:106–114.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe,

Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, and the International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921.

Linnaeus, C. 1758. *Systema Naturae Per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis*. Tomus I, Editio decima, reformata. Laurentius Salvius, Holmiae (Stockholm).

López-García, P., and D. Moreira. 2009. Yet viruses cannot be included in the tree of life. *Nat. Rev. Microbiol.* 7:615−617.

Luo, G., S. Ono, N. J. Beukes, D. T. Wang, S. Xie, and R. E. Summons. 2016. Rapid oxygenation of Earth's atmosphere 2.33 billion years ago. *Sci. Adv.* 2(5):1600134.

Luria, S. E., and J. Darnell. 1967. *General Virology*. John Wiley, New York.

Martel, J., and J. D. E. Young. 2008. Purported nanobacteria in human blood as calcium carbonate nanoparticles. *Proc. Natl. Acad. Sci. USA* 105(14):5549–5554.

Martin, W. F., M. C. Weiss, S. Neukirchen, S. Nelson-Sathi, and F. L. Sousa. 2016. Physiology, phylogeny and LUCA. *Microbial Cell* 3(12):582–587.

Maynard Smith, J., and E. Szathmáry. 1995. *The Major Transitions in Evolution*. W. H. Freeman, New York.

Maynard Smith, J., and E. Szathmáry. 1997. *The Major Transitions in Evolution*. Oxford University Press, Oxford.

McClintock, B. 1950. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* 36(6):344–355.

McDowall, J. 2006. *Protein of the Month: Transposase*. European Bioinformatics Institute. Available at https://www.ebi.ac.uk/interpro/potm/2006_12/Page1.htm.

McKay, D. S., E. K. Gibson, K. L. Thomas-Keprta, H. Vali, C. S. Romanek, S. J. Clemett, X. D. Chillier, C. R. Maechling, and R. N. Zare. 1996. Search for past life on Mars: possible relic biogenic activity in Martian meteorite ALH84001. *Science* 273(5277):924–930.

Mervis, C. B., and E. Rosch. 1981. Categorization of natural objects. *Annu. Rev. Psychol.* 32(1):89–115.

Moreira, D., and P. López-García. 2009. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* 7(4):306–311.

Nakano, Y., H. Aso, A. Soper, E. Yamada, M. Moriwaki, G. Juarez-Fernandez, Y. Koranagi, and K. Sato. 2017. Conflict of interest: the evolutionary arms race between mammalian APOBEC3 and lentiviral Vif. *Retrovirology* 14(31). doi:10.1186/s12977-017-0355-4.

Nasir, A., and G. Caetano-Anollés. 2015. A phylogenomic data-driven exploration of viral origins and evolution. *Sci. Adv.* 1(8):e1500527.

Novikova, N., P. De Boever, S. Poddubko, E. Deshevaya, N. Polikarpov, N. Rakova, L. Coninx, and M. Mergeay. 2006. Survey of environmental biocontamination on board the International Space Station. *Res. Microbiol.* 157(1):5–12.

Orgel, L. E. 1994. The origin of life on the earth. *Sci. Am.* 271(4):76–83.

Oró, J., S. L. Miller, and A. Lazcano. 1990. The origin and early evolution of life on Earth. *Annu. Rev. Earth Planet. Sci.* 18(1):317–356.

Pace, N. R. 2006. Time for a change. *Nature* 441:289.

Pelletier, B. 2015. *Empire Biota: Taxonomy and Evolution*. 1st edition. Createspace Independent Publishing Platform.

Penny, D., and A. Poole. 1999. The nature of the last universal common ancestor. *Curr. Opin. Genet. Dev.* 9:672–677.

Philippe, N., M. Legendre, G. Doutre, Y. Couté, O. Poirot, M. Lescot, D. Arslan, V. Seltzer, L. Bertaux, C. Bruley, J. Garin, J.-M. Claverie, and C. Aberge. 2013. Pandoraviruses: *Amoeba* viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341(6143):281–286.

Podolsky, S. 1996. The role of the virus in origin-of-life theorizing. *J. Hist. Biol.* 29:79–126.

Prestel, E., C. Regeard, J. Andrews, P. Oger, and M. S. DuBow. 2012. A novel bacteriophage morphotype with a ribbon-like structure at the tail extremity. *Res. J. Microbiol.* 7(1):75–81.

Rodrigues-Oliveira, T., A. Belmok, D. Vasconcellos, B. Schuster, and C. M. Kyaw. 2017. Archaeal s-layers: overview and current state of the art. *Front. Microbiol. Microbial Physiol. Metab.* doi.org/10.3389/fmicb.2017.02597.

Root-Bernstein, M., and R. Root-Bernstein. 2015. The ribosome as a missing link in the evolution of life. *J. Theor. Biol.* 367:130–158.

Sára, M., and U. B. Sleytr. 2000. S-layer proteins. *J. Bacteriol.* 182(4):859–868.

Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. D. Reily, L. Courtney, S. S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. M. Rock, E. Belter, F. Du, K. Kim, R. M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. J. Levy, L. McMahan, P. Van Buren, M. W. Vaughn, K. Ying, C. T. Yeh, S. J. Emrich, Y. Jia, A. Kalyanaraman, A. P. Hsia, W. B. Barbazuk, R. S. Baucom, T. P. Brutnell, N. C. Carpita, C. Chaparro, J. M. Chia, J. M. Deragon, J. C. Estill, Y. Fu, J. A. Jeddeloh, Y. Han, H. Lee, P. Li, D. R. Lisch, S. Liu, Z. Liu, D. H. Nagel, M. C. McCann, P. SanMiguel, A. M. Myers, D. Nettleton, J. Nguyen, B. W. Penning, L. Ponnala, K. L. Schneider, D. C. Schwartz, A. Sharma, C. Soderlund, N. M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T. K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. L. Bennetzen, R. K. Dawe, J. Jiang, N. Jiang, G. G. Presting, S. R. Wessler, S. Aluru, R. A. Martienssen, S. W. Clifton, W. R. McCombie, R. A. Wing, and R. K. Wilson2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326(5956):1112–1115.

Sender, R., S. Fuchs, and R. Milo. 2016. Revised estimates for the number of human and bacterial cells in the body. *PLOS Biol.* doi.org/10.1101/036103.

Sereno, P. C. 1998. A rationale for phylogenetic definitions, with application to the higher-level taxonomy of *Dinosauria*. *Neues Jahrb. Geol. Paläontol.* 210:41–83.

Sharp, P. M., and B. H. Hahn. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb. Perspect. Med.* 1:a006841.

Sommer, A. P., D. S. McKay, N. Çiftçioğlu, U. Oron, A. R. Mester, and E. O. Kajander. 2003. Living nanovesicles chemical and physical survival strategies of primordial biosystems. *J. Proteome Res.* 2(4):441–443.

Spinola, M. M. 1850. Tavola sinottica dei generi spettani all classe degli insetti arthroidignati, *Hemiptera* Linn., Latr.—*Rhyngota* Fab.—*Rhynchota* Burm. *Mem. Soc. Ital. Sci. Nat.* 25(1):1–138.

Stejneger, L. 1901. Scharff's history of the European fauna. *Am. Nat.* 35(410):87–116.

Swain, F. M. 1969. Paleomicrobiology. *Annu. Rev. Microbiol.* 23:455–472.

Theobald, D. L. 2010. A formal test of the theory of universal common ancestry. *Nature* 465:219–222.

Torrella, F., and R. Y. Morita. 1981. Microcultural study of bacterial size changes and microcolony and ultramicrocolony formation by heterotrophic bacteria in seawater. *Appl. Environ. Microbiol.* 41(2):518–527.

Trifonov, E. N., and E. Kejnovsky. 2015. *Acytota*—associated kingdom of neglected life. *J. Biomol. Struct. Dyn.* 34(8):1641–1648.

Van Regenmortel, M. H. 2003. Viruses are real, virus species are man-made, taxonomic constructions. *Arch. Virol.* 148:2481–2488.

Villarreal, L. P. 2005. *Viruses and the Evolution of Life*. ASM Press, Washington, DC.

Vorontsov, N. N. 1965. *Origin of Life and Diversity of Its Forms*. Novosibirsk (in Russian).

Wagner, G. P. 2014. *Homology, Genes, and Evolutionary Innovation*. Princeton University Press, Princeton and Oxford.

Wagner, J. R. 2004. The general case of phylogenetic definitions, alternate classes of definitions, and the phylogenetic definition of Life. P. 18 in *Abstracts of the First International Phylogenetic Nomenclature Meeting,* Paris. Available at https://www.phylocode.org and https://hal.sorbonne-universite.fr/hal-02187647.

Walton, L. B. 1930. Studies concerning organisms occurring in water supplies with particular reference to those founded in Ohio. *Ohio Biol. Surv. Bull.* 24:1–86.

Wang, X., N. J. Planavsky, A. Hofmann, E. E. Saupe, B. P. De Corte, P. Philippot, S. V. LaLonde, N. E. Jemison, H. Zou, F. Ossa, K. Rybacki, N. Alfimova, M. J. Larson, H. Tsikos, P. W. Fralick, T. M. Johnson, A. C. Knudsen, C. T. Reinhard, and K. O. Konhauser. 2018. A Mesoarchean shift in uranium isotope systematics. *Geochim. Cosmochim. Acta* 238:438–452.

Wiemann, J., M. Fabbri, T. R. Yang, K. Stein, P. M. Sander, M. A. Norell, and D. E. Briggs. 2018. Fossilization transforms vertebrate hard tissue proteins into N-heterocyclic polymers. *Nat. Commun.* 9(1):4741.

Williams, T. A., P. G. Foster, C. J. Cox, and T. M. Embley. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504(7479):231.

Williams, T. A., P. G. Foster, T. M. W. Nye, C. J. Cox, and T. M. Embley. 2012. A congruent phylogenomic signal places eukaryotes with *Archaea*. *Proc. R. Soc. Lond. B Biol. Sci.* 279:4870–4879.

Woese, C. R. 1999. The universal ancestor. *Proc. Natl. Acad. Sci. USA* 95:6854–6859.

Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc. Natl. Acad. Sci. USA* 87(12):4576–4579.

Wolf, Y. I., D. Kazlauskas, J. Iranzo, A. Lucia-Sanz, J. H. Kuhn, M. Krupovic, V. V. Dolja, and E. V. Koonin. 2018. Origins and evolution of the global RNA virome. *mBio* 9(6): e02329-18.

Yang, S., G. Xin, J. Meng, A. Zhang, Y. Zhou, M. Long, B. Li, W. Deng, L. Jin, S. Zhao, D. Wu, Y. He, C. Li, S. Liu, Y. Huang, H. Zhang, and L. Zou. 2018. Metagenomic analysis of bacteria, fungi, bacteriophages, and helminths in the gut of Giant Pandas. *Front. Microbiol.* 9:1717.

Authors

Jasmina Wiemann; Department of Geology and Geophysics; Yale University; 210 Whitney Avenue; New Haven, CT 06511, USA. Email: jasmina.wiemann@yale.edu.

Kevin de Queiroz; Department of Vertebrate Zoology; National Museum of Natural History, Smithsonian Institution; Washington, DC 20560, USA. Email: dequeirozk@si.edu.

Timothy B. Rowe; Jackson School of Geosciences,;University of Texas at Austin; C1100, Austin, TX 78712, USA. Email: rowe@mail.utexas.edu.

Noah J. Planavsky; Department of Geology and Geophysics; Yale University; 210 Whitney Avenue; New Haven, CT 06511, USA. Email: noah.planavsky@yale.edu.

Ross P. Anderson; Department of Earth Sciences; University of Oxford; Oxford, OX1 3AN, UK. Email: ross.anderson@all-souls.ox.ac.uk.

Johann P. Gogarten; Department of Molecular and Cell Biology; University of Connecticut; Storrs, CT 06269 USA. Email: gogarten@uconn.edu.

Paul E. Turner; Department of Ecology and Evolutionary Biology; Yale University; P. O. Box 208106, New Haven, CT 06520, USA. Email: paul.turner@yale.edu.

Jacques A. Gauthier; Department of Geology and Geophysics; Yale University; 210 Whitney Avenue; New Haven, CT 06511, USA. Email: jacques.gauthier@yale.edu.

Date Accepted: 11 July 2019

Primary Editor: Philip Cantino