A Genomic Perspective Across Earth's Microbiomes Reveals That Genome Size in Archaea and Bacteria Is Linked to Ecosystem Type and Trophic Strategy

Alejandro Rodríguez-Gijón1†, Julia K. Nuy1†, Maliheh Mehrshad2, Moritz Buck2, Frederik Schulz3, Tanja Woyke3 and Sarahi L. Garcia1*

1 Department of Ecology, Environment, and Plant Sciences, Science for Life Laboratory, Stockholm University, Stockholm, Sweden, 2 Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden, 3 DOE Joint Genome Institute, Berkeley, CA, United States

Our view of genome size in Archaea and Bacteria has remained skewed as the data has been dominated by genomes of microorganisms that have been cultivated under laboratory settings. However, the continuous effort to catalog Earth’s microbiomes, specifically propelled by recent extensive work on uncultivated microorganisms, provides an opportunity to revise our perspective on genome size distribution. We present a meta-analysis that includes 26,101 representative genomes from 3 published genomic databases; metagenomic assembled genomes (MAGs) from GEMs and stratfreshDB, and isolates from GTDB. Aquatic and host-associated microbial genomes present on average the smallest estimated genome sizes (3.1 and 3.0 Mbp, respectively). These are followed by terrestrial microbial genomes (average 3.7 Mbp), and genomes from isolated microorganisms (average 4.3 Mbp). On the one hand, aquatic and host-associated ecosystems present smaller genomes sizes in genera of phyla with genome sizes above 3 Mbp. On the other hand, estimated genome size in phyla with genomes under 3 Mbp showed no difference between ecosystems. Moreover, we observed that when using 95% average nucleotide identity (ANI) as an estimator for genetic units, only 3% of MAGs cluster together with genomes from isolated microorganisms. Although there are potential methodological limitations when assembling and binning MAGs, we found that in genome clusters containing both environmental MAGs and isolate genomes, MAGs were estimated only an average 3.7% smaller than isolate genomes. Even when assembly and binning methods introduce biases, estimated genome size of MAGs and isolates are very similar. Finally, to better understand the ecological drivers of genome size, we discuss on the known and the overlooked factors that influence genome size in different ecosystems, phylogenetic groups, and trophic strategies.

Keywords: microbial ecology, genome size, bacteria, archaea, genomics
INTRODUCTION

As microbiologists, how do we define what is a small or a big genome? Perhaps, researchers working on model organisms such as *Escherichia coli* with a genome size of ~5 Mbp (Abram et al., 2021) would define “big” or “small” differently to researchers working on soil-dwelling bacteria with a genome size of 16 Mbp (Garcia et al., 2014). On the lower genome size scale, whereas genome sizes of bacterial endosymbionts of insects may have genomes merely larger than 100 kbp (Moran and Bennett, 2014), the abundant *Prochlorococcus* range between 1.6 and 1.9 Mbp for high-light and low-light ecotypes (Berube et al., 2018). In summary, it is known that genome sizes of Archaea and Bacteria range between 100 kbp and 16 Mbp, but the genome size distribution in nature is still undefined. Therefore, the aim of this review is to provide an overview of the distribution of genome sizes in different ecosystems.

We leveraged recently published databases of archaeal and bacterial metagenome assembled genomes (MAGs) (Nayfach et al., 2020; Buck et al., 2021a) together with isolate genomes to revisit and acquire an updated understanding of the estimated genome size distribution across different ecosystems. In this review, we also discuss the ecological drivers that potentially influence genome sizes. In summary, we found that 76.3% of representative archaecal and bacterial genomes recovered through genome-resolved metagenomics present estimated genome sizes below 4 Mbp. Furthermore, all MAGs from five archaecal phyla (Micarchaeota, Ianarchaeota, Undinarchaeota, Nanohaloarchaeota, and Hadarchaeota) and two bacterial phyla (Coprothermobacterota and Dictyoglomota) were recovered exclusively from aquatic ecosystems and have genome sizes below 2 Mbp (Figures 1A,B).

APPROXIMATION OF GENETIC UNITS USING 95% AVERAGE NUCLEOTIDE IDENTITY AND ITS CAVEATS

Species are widely considered congruent genetic and ecological units for sexual eukaryotes (Mallet, 2008; Shapiro and Polz, 2015). However, there is no consensus regarding the concept of species for Archaea and Bacteria. Instead, 95% average nucleotide identity (ANI) has been a widely recognized as a genetic boundary to operationally estimate genetic units or “microbial species” (Konstantinidis and Tiedje, 2005; Varghese et al., 2015; Garcia et al., 2018; Jain et al., 2018). Several genomic and metagenomic studies have verified the existence of sequence discrete genetic units with 95% ANI as boundary (Olmi et al., 2020; Rodriguez et al., 2021).

We used the 95% ANI boundary in published datasets (Nayfach et al., 2020; Parks et al., 2020; Buck et al., 2021a) to review and renew our view of the distribution of archaecal and bacterial estimated genome size. To minimize representation biases (Gweon et al., 2017), we included one representative per 95% ANI cluster. The approximation of these clusters will be called mOTUs (metagenomic Operational Taxonomic Units) (Buck et al., 2021b; Garcia et al., 2021). We used only one MAG per mOTU, making the estimated genome size dependent on the MAG assembly and quality. It is known that MAG assembly and binning might discriminate against ribosomal RNAs, transfer RNAs, mobile element functions and genes of unknown function (Nelson et al., 2020; Meziti et al., 2021), and also that completeness could be underestimated for streamlined genomes (Garcia et al., 2015). Despite potential biases introduced by these methodological limitations, we still offer valuable insights on genome size distribution across different environments.

EXTANT GENOME SIZE DISTRIBUTION IN THE ENVIRONMENT

In this review, we have included ~64,500 environmental MAGs available via two recently published datasets, stratfreshDB and GEMs. StratfreshDB offers ~12,000 MAGs (>40% completeness) from 41 stratified lakes and ponds assembled with Megahit (v1.1.13) and binned with Metabat (v2.12.1) (Buck et al., 2021a). GEMs offers ~52,000 MAGs (>50% completeness) from >10,000 metagenomes collected from diverse habitats on Earth (Nayfach et al., 2020). GEMs dataset was assembled using metaSPAdes and binned with Metabat (v0.32.5). After dereplication using fastANI and mOTUlizer (Jain et al., 2018; Buck et al., 2021b), our meta-analysis includes 17,834 mOTUs, with one representative MAG each (completeness >50% and contamination <5%, assessed with CheckM v1.1.3) (Parks et al., 2015). We complemented these environmental MAGs by adding 8,267 representative genomes (>90% complete) of isolates from GTDB (Parks et al., 2020; Figure 1A). These genomes are marked in the GTDB databases (release 95) to originate from culture collections. After clustering at 95% ANI threshold, 540 mOTUs contained representatives from both environmental MAGs and isolate genomes (Figure 1C). Previous surveys based on 16S rRNA have found that the uncultured microbial fraction could constitute up to 81% of the total microbial diversity on Earth (Lloyd et al., 2018). However, it is known that 16S rRNA underestimate prokaryotic diversity (Rodriguez et al., 2018). Overall, our review shows that 3% of the reconstructed environmental mOTUs are represented among cultured microbes.

Furthermore, using completeness estimates from CheckM, we compared the estimated genome size distribution of all MAGs vs. genomes from isolates. The estimated genome size was calculated by dividing the MAG assembly size by CheckM completeness (ranging from 0 to 1). Representative genomes from isolates have an average genome size of 4.3 Mbp which is significantly larger than that of MAGs (t-test p < 0.0001), both when comparing Archaea and Bacteria combined and separately. To compare estimated genome sizes between MAGs, ecosystem type was used according to the GEMs database. Although the ecosystem classification presented here is coarse and might contain countless niches, it still allowed us to see trends for genome sizes. Estimated genome sizes of aquatic MAGs have an average of 3.1 Mbp, host-associated MAGs average 3.0 Mbp, and terrestrial MAGs average 3.7 Mbp (Figure 1A). For the 540 mOTUs that contained both environmental MAGs and isolate
FIGURE 1 | Overview of the genome size distribution across Earth’s microbiomes. Genome size distribution of Archaea and Bacteria (A) from different environmental sources and across different archaeal and bacterial phyla (B) are shown for a total of 26,101 representative genomes. Isolate genomes were gathered from GTDB (release95) and environmental MAGs were gathered from GEMs (Nayfach et al., 2020) and stratfreshDB (Buck et al., 2021a). We use one representative genome per mOTU (defined by 95% ANI) from the union of GEMs catalog and stratfreshDB in the plots. From the GTDB database, we selected one representative isolate genome per species cluster that was circumscribed based on the ANI (>95%) and alignment fraction (AF) > 65% between genomes (Parks et al., 2020). To construct the figures, we plotted the min-max estimated genome sizes, which were calculated based on the genome assembly size and completeness estimation provided. Venn diagram of the intersection between the representative environmental MAGs and the representative isolate genomes (C). The intersection was calculated using FastANI (Jain et al., 2018) and was determined with a threshold of 95%. The coding density (D) and GC content (%) (E) are shown for the archaeal and bacterial MAGs across different ecosystem categories and isolates. Pair-wise t-test was performed in all variables of (D,E) and shown in (F), where white is significant (p < 0.05) and black is not significant (p > 0.05). In (B), we only included phyla with more than five genomes.
In each niche, there will be different selective pressures on the genome size. An example is clearly shown in a study (Nayfach and Pollard, 2015) in which it is observed that Archaea and Bacteria sampled from different parts of the human body have differences in genome sizes. Low metadata resolution and clustering of all genomes into three main ecosystem types might be a reason why we see a range of genome sizes in the genera of different ecosystems (Figure 2). With more precise metadata and higher sampling resolution of microhabitats, it might be possible to identify the ecological drivers of genome sizes in the different niches in fine-scale.

**IMPACT OF ECOSYSTEM AND TROPHIC STRATEGY ON GENOME SIZE**

Terrestrial ecosystems harbor immense microbial diversity (Delgado-Baquerizo et al., 2018). Yet, the most up-to-date data compilation provided here shows only 2033 MAGs from terrestrial ecosystems (Figure 1C) with an average genome size of 3.7 Mb (Figure 1A). The sub-ecosystems considered in this view are soil and deep subsurface, among others (Supplementary Figure 2). While the terrestrial microorganism’s genome size is the biggest of the three ecosystem categories in this review, they are smaller than expected based on previous metagenomic predictions, which placed the genome size of soil bacteria at 4.74 Mb (Raes et al., 2007). Trends for larger genome sizes in soil have been hypothesized to be related to scarcity and high diversity of nutrients, and a fluctuating environment combined with little penalty for the slow growth rate (Konstantinidis and Tiedje, 2004; Cobo-Simon and Tamames, 2017; Chen et al., 2021).

Although terrestrial environments are physically structured, they are generally characterized by two to three orders of magnitude greater variations (in temperature and currents) than marine environments (Steele et al., 2019). In silico studies predict that large genome sizes could result from higher environmental variability (Bentkowski et al., 2015). A recent example showed that isolates of terrestrial Cyanobacteria have genomes on the larger size scale (6.0–8.0 Mb) that are enriched in genes involved in regulatory, transport and motility functions (Chen et al., 2021). These functional categories enable them to thrive in a fluctuating environment with high nutrient diversity. Despite these general trends showing larger genome sizes in terrestrial ecosystems, it is worth noting that the diversity captured in the GEMs survey is probably a small fraction of the total terrestrial microbial diversity. It is, for example, also known that streamlined microorganisms such as Patescibacteria (Ortiz et al., 2021) and “Candidatus Udaeobacter copiosus” (Verrucomicrobiota) are abundant in soils (Brewer et al., 2016). Therefore, we predict that the view on genome size distribution and microbial diversity in terrestrial ecosystems will become more complete with more sequencing, assembly, binning and novel isolation efforts.

In host-associated microbiomes, genetic drift, deletion biases and low populations sizes drive the reduction of genomes (Li et al., 2021). In these environments, the differing levels of intimacy with their host can influence the evolution of the genome of microorganisms. For example, within the
FIGURE 2 | Phylogenetic trees of archaeal (A) and bacterial (B) representative genomes show variation in genome size between and within phyla. The trees were constructed using GTDB-tk (v 1.5.0) using de novo workflow using aligned concatenated set of 122 and 120 single copy marker proteins for Archaea and Bacteria, respectively (Chaumeil et al., 2020). Moreover, in this mode, GTDB-tk adds 1,672 and 30,238 backbone genomes for Archaea and Bacteria, respectively. Tree is visualized in anv’i’o (Eren et al., 2021). Estimated genome size is presented in scale from 0 to 6 Mbp or 0 to14 Mbp for Archaea or Bacteria. In the tree, the origin of the environmental genomes is labeled: aquatic, terrestrial and host-associated (same MAGs as Figure 1). Highlighted phyla with more representative genomes are color-coded. Boxplots show the average estimated genome size per phyla within archaeal and bacterial (C) domains. The average estimated size per genus within Halobacteriota (D), Thermoproteota (E), Actinobacteriota (F), Bacteriodiota (G), Firmicutes A (H), Patescibacteria (I), Proteobacteria (J). The presence of phyla and genera is colored in gray if they contain MAGs from different ecosystem category (non-specific ecosystem). The average estimated size per genus extracted from aquatic ecosystems (K), host-associated ecosystems (L), terrestrial ecosystems (M), or non-specific ecosystems (N). Letters in boxplot panels are the result of non-parametric tests, Wilcoxon (C) or Kruskal-Wallis (D–N). Different letters show significant differences p < 0.05 (all statistical test with multiple testing were corrected with Benjamini–Hochberg).
of host-associated MAGs show coding densities of ∼91% for genomes below 2 Mbp (Figure 1D).

In this review, the largest fraction of MAGs is recovered from aquatic environments. The two main sub-ecosystems in our survey are freshwater with MAGs estimated average genome size of 3.2 Mbp significantly different (p < 0.0001) from marine genome size distribution with average estimated genome size of 2.9 Mbp. When comparing freshwater and marine environments, the most obvious difference is salinity followed by nutrient concentration. Further exploring the impact of differing levels of salinity on genome size is an interesting research prospect. Additionally, we compared the union of representative freshwater MAGs from both databases (StratfreshDB and GEMs) (Supplementary Figure 3). The difference of mean estimated genome size between the representatives from freshwater GEMs and StratfreshDB is 0.52 Mbp. However, this is because each database captures genetic units that were not found in the other database.

In general, aquatic environments are vertically structured by gradients of light penetration, temperature, oxygen, and nutrient (Supplementary Table 1). Moreover, microorganisms might experience a microscale spatial and nutrient structure due to the presence of heterogeneous particles. These aquatic structures are drivers of the genetic repertoire of aquatic microorganisms. Metagenomic sequencing reported the increase of genome sizes for Archaea and Bacteria with increasing depths (Mende et al., 2017). Temperature may be as important; for example, a study based on twenty-one Thermoproteota and Euryarchaeota fosmids (Euryarchaeota currently reclassified into Methanobacteriota, Halobacteriota, and Nanohaloarchaeota) showed high rates of gene gains through HGT to adapt to cold and deep marine environments (Brochier-Armanet et al., 2011). It has been observed that light is a relevant driver of genome size in aquatic environments as it decreases with depth. Photosynthetic bacteria such as Prochlorococcus spp. are well-differentiated into a high-light adapted ecotype with smaller genome sizes (average 1.6 Mbp) and a low-light-adapted ecotype with a slightly bigger genome size (average 1.9 Mbp) (Berube et al., 2018; Supplementary Table 1). Limitation of nutrients such as nitrogen (Elser et al., 2007) might also be one of the central factors determining genomic properties (Grzymski and Dussaqu, 2012). Nitrogen fixation is a complex process that requires a great amount of genes (Franche et al., 2008) and most nitrogen-fixing marine cyanobacteria have the largest genomes in its phylum (Bergman et al., 2013).

Diversity and quantity of nutrients might be two understudied factors that drive ecology and genome size evolution. A recent example shows that polysaccharide xylan triggers microcolonies, whereas monosaccharide xylose promotes solitary growth in Caulobacter (D’Souza et al., 2021). This is a striking example of how nutrient complexity can foster diverse niches for well-studied cells such as Caulobacter with genome size 4 Mbp. To fully understand the link between genome size and nutritional requirements of diverse environmental microorganisms, we need to systematically explore the ∼90% of molecules/metabolites still unknown (Wienhausen et al., 2017; Hawkes et al., 2018; Patriarca et al., 2020). The wide nutrient complexity in the environment might prompt microorganisms to shape their genome. Their genomic content and metabolic potential defines whether they
are capable to feed on the available nutrients, forcing them to develop dependencies with other community members in order to acquire energy and metabolic precursors. Metagenomics combined with metabolomics will provide an understanding of the link between genome size evolution of microorganisms and their nutritional and trophic strategy.

CONCLUSION

This review offers a broad overview of genome size distribution across three different ecosystem categories, showing that MAGs recovered from aquatic and host-associated ecosystems present smaller estimated genome sizes than those recovered from terrestrial ecosystems. Moreover, genomes obtained from environmental samples present a smaller estimated genome size than obtained by cultivation approaches. We find that the distribution of genome sizes across the phylogenetic tree of Archaea and Bacteria can be linked to the ecosystem type from which the microorganisms’ genomes have been extracted (aquatic, host-associated or terrestrial). Finally, we review the ecological factors that may cause the varying sizes of genomes in different ecosystems. In comparison with the aquatic and host-associated ecosystems, terrestrial ecosystems might harbor microorganisms with bigger estimated genome sizes mainly due to higher fluctuations in this ecosystem. Host-associations might shape genomes sizes differentially based on the type of host and level of intimacy between the microorganisms and the host. Genomes in aquatic ecosystems might be shaped by vertical stratification of abiotic factors such as nutrient distribution, light penetration, and temperature. Moreover, different trophic strategies such as autotrophs might be connected to smaller genome sizes. We expect that as the microbial ecology field keeps moving forward with sequencing, bioinformatics, chemical analysis, and novel cultivation techniques, we will get a deeper resolution on physicochemical, metabolic, spatial, and biological drivers of archaecal and bacterial genome sizes.

REFERENCES

Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T. M., Robeson, M. S. II, et al. (2021). Mass-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups. Commun. Biol. 4:117. doi: 10.1038/s42255-020-01626-5
Aylward, F. O., and Santoro, A. E. (2020). Heterotrophic Thaumarchaea with small genomes are widespread in the dark ocean. mSystems 5, e00415–e00420. doi: 10.1128/mSystems.00415–20
Bentkowski, P., Van Oosterhout, C., and Mock, T. (2015). A model of genome size evolution for prokaryotes in stable and fluctuating environments. Genome Biol. Evol. 7, 2344–2351. doi: 10.1093/gbe/evv148
Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E. J. (2013). Trichodesmium—a widespread marine cyanobacterium with unusual nitrogen fixation properties. FEMS Microbiol. Rev. 37, 286–302. doi: 10.1111/j.1574-6976.2012.00352.x
Berube, P. M., Biller, S. J., Hackl, T., Hogle, S. L., Satinsky, B. M., Becker, J. W., et al. (2018). Single cell genomes of Prochlorococcus. Synechococcus, and sympatric microbes from diverse marine environments. Sci. Data 5:180154. doi: 10.1038/sdata.2018.154
Boussau, B., Karlberg, E. O., Frank, A. C., Legault, B. A., and Andersson, S. G. (2004). Computational inference of scenarios for alpha-proteobacterial genome evolution. Proc. Natl. Acad. Sci. U.S.A. 101, 9722–9727. doi: 10.1073/pnas.0400975101
Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., and Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium 'Candidatus Udaeobacter copiosus'. Nat. Microbiol. 2:16198. doi: 10.1038/nmicrobiol.2016.198
Brochier-Armanet, C., Deschamps, P., Lopez-Garcia, P., Zivanovic, Y., Rodriguez-Valera, F., and Moreira, D. (2011). Complete fosmid and fosmid-end sequences reveal frequent horizontal gene transfers in marine uncultured planktonic archaea. ISME J. 5, 1291–1302. doi: 10.1038/ismej.2011.16
Buck, M., Garcia, S. L., Fernandez, L., Martin, G., Martinez-Rodriguez, G. A., Saarenheimo, J., et al. (2021a). Comprehensive dataset of shotgun metagenomes from oxygen stratified freshwater lakes and ponds. Sci. Data 8:131. doi: 10.1038/s41597-021-00910-1
Buck, M., Mehrshad, M., and Bertilsson, S. (2021b). mOTUpan: a robust Bayesian approach to leverage metagenome assembled genomes for core-genome estimation. bioRxiv [Preprint]. doi: 10.1101/2021.06.25.449606
Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S. J. (2013). Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a defined medium. ISME J. 7, 592–602. doi: 10.1038/ismej.2012.122

AUTHOR CONTRIBUTIONS

SG, AR-G, and JN conceptualized the literature and data review idea. JN, MB, and FS gathered the data. AR-G and JN performed data analysis. SG, AR-G, and MM drafted the first manuscript. All authors did literature searches, contributed to the writing, and editing of the manuscript.

FUNDING

This work was supported by SciLifeLab and Kungl. Vetenskapsakademien's stifelser grant CR2019-0060. The computations and data handling were enabled by resources in the project SNIC 2021/6-99 and SNIC 2021/5-133 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The work conducted by the U.S. Department of Energy Joint Genome Institute, an Office of Science User Facility, made use of the National Energy Research Scientific Computing Center and was supported under Contract No. DE-AC02-05CH11231.

ACKNOWLEDGMENTS

We are grateful to John Paul Balmonte, Sergio Tusso, and Alexander Probst for helpful discussions. AR-G thanks Fede Berckx for technical advice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.761869/full#supplementary-material
Nayfach, S., and Pollard, K. S. (2015). Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol.* 16:51. doi: 10.1186/s13059-015-0611-7

Nayfach, S., Roux, S., Seshadri, U., Udvardy, D., Varghese, N., Schulz, F., et al. (2020). A genomic catalog of Earth’s microorganisms. *Nat. Biotechnol.* 39, 499–509. doi: 10.1038/s41587-020-0718-6

Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *PeerJ* 8:e10119.

Olm, M. R., Crits-Christoph, A., Diamond, S., Lavy, A., Matheus Carnevali, P. B., Patriarca, C., Sedano-Núñez, V. T., Garcia, S. L., Bergquist, J., Bertilsson, S., Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *Genome Biol.* 8:e10119.

Olm, M. R., Crits-Christoph, A., Diamond, S., Lavy, A., Matheus Carnevali, P. B., Patriarca, C., Sedano-Núñez, V. T., Garcia, S. L., Bergquist, J., Bertilsson, S., Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *PeerJ* 8:e10119.

Olm, M. R., Crits-Christoph, A., Diamond, S., Lavy, A., Matheus Carnevali, P. B., Patriarca, C., Sedano-Núñez, V. T., Garcia, S. L., Bergquist, J., Bertilsson, S., Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *PeerJ* 8:e10119.

Parks, D. H., C. L., Kan, B., Lavy, A., Matheus Carnevali, P. B., Patriarca, C., Sedano-Núñez, V. T., Garcia, S. L., Bergquist, J., Bertilsson, S., Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *PeerJ* 8:e10119.

Perspect. Biol. 7:a0181431.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rodríguez-Gijón, Nuy, Mehrshad, Back, Schulz, Woyke and García. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.