Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Solar salterns as model systems to study the units of bacterial diversity that matter for ecosystem functioning

Konstantinos T Konstantinidis[1,2,3], Tomeu Viver[4], Roth E Conrad[1,2], Stephanus N Venter[5] and Ramon Rossello-Mora[4]

Microbial communities often harbor overwhelming species and gene diversity, making it challenging to determine the important units to study this diversity. We argue that the reduced, and thus tractable, microbial diversity of manmade salterns provides an ideal system to advance this cornerstone issue. We review recent time-series genomic and metagenomic studies of the saltern-dominating bacterial and archaeal taxa to show that these taxa form persistent, sequence-discrete, species-like populations. While these populations harbor extensive intra-population gene diversity, even within a single saltern site, only a small minority of these genes appear to be functionally important during environmental perturbations. We outline an approach to detect and track such populations and their ecologically important genes that should be broadly applicable.

Addresses
[1] Ocean Science & Engineering, School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[2] School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[3] School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[4] Marine Microbiology Group, Department of Animal and Microbial Biodiversity, Mediterranean Institutes for Advanced Studies (IMEDEA, CSIC-UIB), Esporles, Spain
[5] Department of Biochemistry, Genetics and Microbiology, and Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

Corresponding authors:
Konstantinidis, Konstantinos T (kostas@ce.gatech.edu), Rossello-Mora, Ramon (rossello-mora@uib.es)

Current Opinion in Biotechnology 2021, **73**:151–157

This review comes from a themed issue on **Environmental biotechnology**

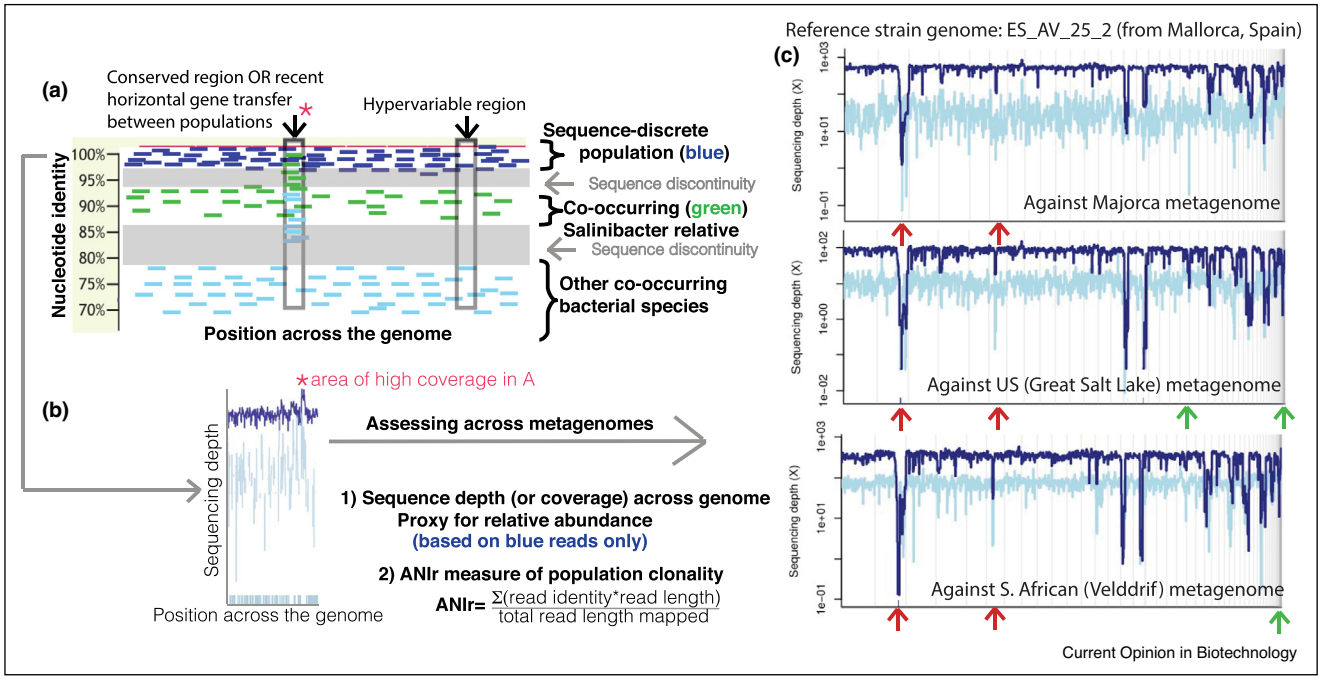
Edited by **Luigi Vezzulli** and **Marco Ventura**

https://doi.org/10.1016/j.copbio.2021.07.028

Natural and engineered microbial communities are often composed of hundreds, if not thousands of species, each carrying a couple hundred species-specific genes [1,2]. Given this high complexity, it has been challenging to decipher which specific taxa and genes carry out the key ecosystem functions. Complicating the matter further, it has been argued that prokaryotes, and other microbes, do not form distinct populations due to the extensive horizontal gene transfer they can undergo [3] and their very large population sizes (e.g. population diversity sweep events are uncommon). Indeed, whether or not prokaryotic species exist and, if yes, how to recognize them are challenging questions to answer, with obvious practical implications in identifying or regulating organisms of clinical or biotechnological importance [1,4,5]. Accordingly, what are the appropriate units to study bacterial diversity within an (eco-)system remains an unresolved issue, despite its paramount importance for research activities as well as communication of results. Consequently, several scientists study operational taxonomy units (OTUs), defined by an arbitrary threshold of nucleotide identity, usually 97% or 99% for the 16S rRNA (or another) gene, as a proxy for a species or unit, while others opt to study individual genes, without considering the genomic background that the genes reside in.

During the previous decade, shotgun metagenomics studies of natural microbial populations have revealed that bacteria and archaea predominantly form sequence-discrete populations with intra-population genomic sequence relatedness typically ranging from ~95% to ~100% genome-aggregate average nucleotide identity (or ANI) depending on the population considered (e.g. younger populations since the last population diversity sweep event show lower levels of intra-population diversity). In contrast, ANI values between distinct populations are typically lower than 90% [6] (Figure 1a). Such sequence-discrete populations were recovered from many different habitats, including fresh-waters and marine-waters, soils, sediments, human gut microbiomes and biofilms of engineered systems, and were typically persistent over time and space [7–9,10•] indicating that they are not ephemeral but long-lived entities. The qualitative difference between these metagenomics results and those of previous studies based on isolated organisms in the laboratory that reported indiscrete species may be due to isolation biases [6]. That is, the previous studies included heterogeneous collections of organisms representing different locations and sampling times (i.e. distinct ecological niches) and genomic adaptations specific to local

Figure 1



The read-recruitment approach to identify and track a sequence-discrete population.
Panel **(a)**: Recruitment plot of metagenomic reads from a control saltern sample against the genomic sequence of a *Sal. ruber* isolate from Mallorca, Spain (reference). Note that the recruited reads representing the same population as the reference genome (in dark blue) vary in identity from ~96 to 100%, indicating substantial sequence and gene content diversity (e.g. denoted by fewer reads mapping in the hypervariable region) within the population (not clonal). Reads representing co-occurring but distinct related populations typically show <95% nucleotide identity to the reference (in green, for the low abundance *Sal. ruber*-like population; in light blue for the rest; see text for further details). Panel **(b)**: Based on the nucleotide identity of the dark blue reads against the reference, one can calculate the ANIr value (formula at the bottom), a metric of the clonality of the population [10•], and the sequencing depth (or coverage) by mapped reads across the reference genome (graph on the bottom, left), the mean of which could be used as a proxy of the relative abundance of the reference in the metagenome [28•]. Panel **(c)**: Coverage of *Sal. ruber* genes by the Spain (Mallorca), US (Great Salt Lake) and South African (Velddrif) saltern metagenomes. The metagenomes were searched against the whole-genome of *Sal. ruber* isolate ES_AV_25_2 from Mallorca (Spain) as shown in panel A. The graphs show the sequencing depth (coverage; y-axis) of the genome (x-axis) by reads representing the *Sal. ruber* (dark blue) and the remaining populations (light blue) present in the metagenomes [thus, coverage in dark blue is basically produced by the dark blue reads of panel A and is similar to that shown in Panel (b)]. Note that several genomic islands, denoted by drops in coverage, are specific to the isolate (red arrows) while others are present in the Mallorca (Spain) but not in the US or S. African *Sal. ruber* populations (green arrows), revealing a core *Sal. ruber* gene set across salterns from different continents but also important gene-content differences in non-core genes. Similar plots were obtained with other Mallorca isolate genomes, revealing that our isolates represent well the *Sal. ruber* populations found in other continents. Reads representing the *Sal. ruber*-like population (in green in Panel (a)) were removed from Panels (b) and (c) for simplicity. The example recruitment plots were obtaining using the genome and metagenome data available through the ENA project PRJEB45291.

environmental conditions at the place of isolation which may have confounded the existence of sequence-discrete, species-like units [11]. It is important to note, however, that a recent analysis of all available isolate genomes ($n = \sim 90\,000$) also revealed sequence discontinuities between most named species around 85–95% ANI [12•] consistent with the picture emerging from metagenomics (but not the early picture from comparisons of isolate genomes). These results constitute a paradigm shift since several scientists believed that microbes do not form distinct populations or species (discussed above). More recent work has even shown that intermediate identity genotypes, for example, sharing 85–95% ANI, when present, are ecologically differentiated and thus, should be probably considered distinct species [13••,14••,15••]. Therefore, the 95% ANI threshold appears to be robust [15••]. Accordingly, we recommend the use of 95% ANI to demarcate the important units of microbial communities, in general, and the methods outlined by the studies mentioned above, and shown in Figure 1, provide the means to identify and track the sequence-discrete populations. Note, however, that we do not claim that the 95% ANI is a hard threshold. On the contrary, we observe this sequence boundary in nature, likely as an emerging property of many eco-evolutionary forces at play, and that this threshold may slightly differ for different taxa (e.g. range between 90 and ~98% ANI). Notably, the ANI-based approach also circumvents the major limitation of 16S rRNA gene-based approaches. That is, closely related, yet distinct populations (and

species) at the genomic level often share >99% 16S rRNA gene identity due to the high sequence conservation of this gene relative to the whole genome and thus, are lumped together by the most commonly used thresholds for defining 16S rRNA gene-based OTUs.

Another major conclusion from genomic studies during the previous decade is that the gene-content diversity within a species is often overwhelming. In other words, the pangenomes of bacterial species, that is, the number of non-redundant genes carried by members of the species, can be enormous based on the genome sequencing of isolates from various sites around the globe and different years. However, the value of this gene diversity for ecosystem functioning has remained challenging to quantify. For instance, it is not clear what fraction of this gene content diversity is ecologically important (non-neutral) as opposed to transient or ephemeral [13••,16,17]. It is also important to note that the diversity within the sequence-discrete populations cannot be efficiently studied solely with shotgun metagenomic methods because advantageous genes (if any) could be at low levels (e.g. have not yet differentiated enough in abundance compared to the pre-selection/pre-transition period to be detectable by read-recruitment plots or other means) and/or show high sequence identity at this (intra-population) level, which prevents — for instance — robust assembly of individual genomes and assessment of gene exchange. The typical read lengths obtained by short-read metagenome sequencing (e.g. 100−500 bp) cannot recover the full length of the average gene (~1000 bp long). Our team has recently obtained novel, quantitative insights into these questions by studying salterns.

Solar salterns are human-controlled, semi-artificial environments used for the harvesting of salt for human consumption. These environments are operated in repeated cycles of feeding with natural saltwater, increasing salt concentration due to water evaporation, and finally, salt precipitation. Several studies have shown that salterns in different parts of the world harbor recurrent microbial communities each year [18,19]. These communities show low, and thus tractable, diversity, generally consisting of two major lineages, that is, the archaeal *Halobacteria* class and the bacterial family of *Salinibacteraceae*, class *Rhodothermia* [19–21], but with relatively high species richness within each class [14••,22••]. To cope with the extreme salt concentrations close to or above salt saturation (i.e. ~36% NaCl), halophilic microorganisms have evolved different osmotic survival strategies such as production of osmoprotectants and compatible solutes for the moderate halophiles and a salt-in strategy (concentrating salts, especially potassium, in the cytoplasm) for the extreme halophiles [23]. Besides salinity, irradiation (sunlight) is probably the second most relevant environmental driver in these systems, and saltern-dwelling microorganisms have distinct DNA repair systems and photolyases to cope with UV radiation stresses [24]. Importantly, a large fraction of the microbial communities in salterns can be readily cultured [14••,25], and it is relatively easy to experimentally manipulate the two main drivers (light intensity and salt concentration) of these systems *in situ* (Figure 2) [22••,26••].

Specifically, we sequenced 110 *Salinibacter ruber* isolate genomes chosen from a larger collection of 220 isolates based on different PCR random amplified polymorphic DNA (RAPD) profiles in order to avoid sequencing the same clone, from a single saltern site in the Mallorca island, Spain. Mapping of metagenomic reads from the samples of origin against these genome sequences revealed that our isolates represent well the natural (metagenomic) population, for example, few new alleles of *Sal. ruber* housekeeping genes, defined as showing <99.9% nucleotide identity to an existing allele, were found among the mapped reads compared to the alleles present in the genomes (our unpublished data). Notably, our isolates seem to represent well the global *Sal. ruber* populations found in other salterns and continents ([13••] and Figure 1c). Thus, our isolate collection constitutes a unique resource to provide quantitative insights into the abovementioned questions. Comparisons among these isolate genomes revealed an intriguing population structure wherein the majority (100/110) showed >97% ANI to each other, with the remaining 10 genomes showing ~93% ANI to the dominant group and >97% ANI among themselves. Hence, the total *Salinibacter* population appears to be composed of a dominant sequence-discrete population and a minor *Sal. ruber*-like population. The latter population is apparently ecologically differentiated from the former dominant population since under the same conditions (e.g. same sample) it shows lower relative abundance and is genetically distinct (Figure 1a), albeit its exact ecological niche remains speculative. Therefore, these two populations further corroborated the applicability of the 95% ANI threshold as outlined above and represent, in addition, a unique system to study population maintenance and diversification in the future.

Whole-genome comparisons revealed that the pangenome of the 100 isolate genomes representing the dominant population *Sal. ruber* composed of about 13 000 non-redundant genes, revealing extensive gene-content diversity that is similar to that observed among 100 randomly sampled *Escherichia coli* genomes from NCBI [13••]. While the *E. coli* genomes originated from different ecological niches and samples, including disease (pathogens), healthy (commensal) and environmental samples, and thus, as noted previously [16,27] an open pangenome was expected, all *Sal. ruber* isolates originated from the same site and time (summer of 2012) and were apparently present in the (same) source of preconcentrated brines. Hence, these results are somewhat surprising, and do

Figure 2

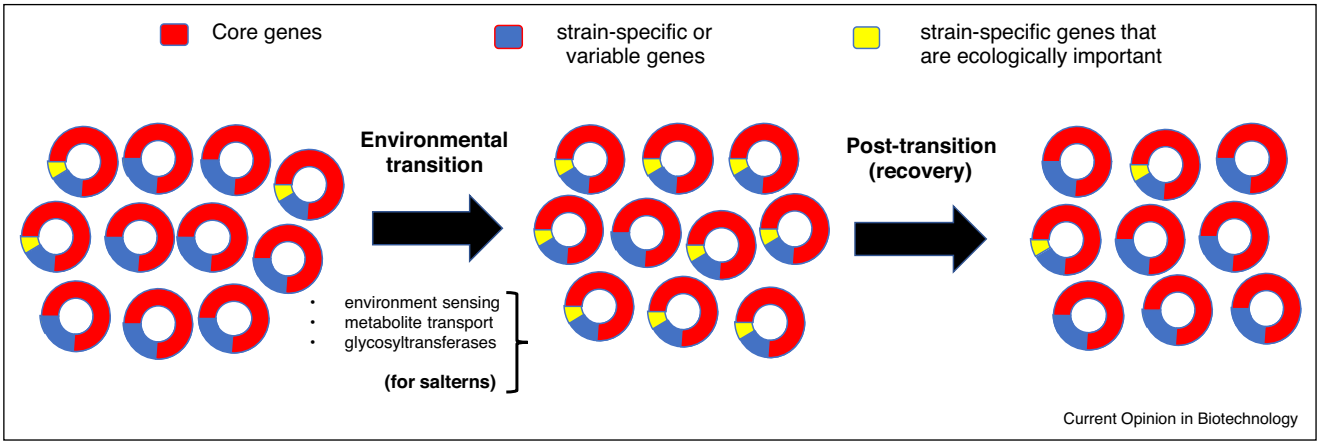
Shading a saltern pond to assess the effect of light intensity in Majorca, Spain.
A pond is shaded with a mesh to assess the effects of sunlight on the indigenous microbial communities (upper) and is located next to a saturated (bottom) pond in terms of salts.

suggest that natural populations like *Sal. ruber* could contain immense gene-content diversity and open pangenomes.

Examination of the relative abundance of the pangenome genes within the *Salinibacter* population in salterns under shade (by the application of a mesh [22••]) and low-salinity (after dilution with fresh-water [26••]) relative to the control (ambient sunlight and salt saturation conditions) revealed interesting patterns. While the majority of genome-specific (i.e. found only in one of the total 100 genomes) and rare genes (i.e. found in <10% of the genomes) remained low-abundance when conditions changed, at least a handful of genes (about 2% of the total) were found to significantly increase in abundance in the low-salinity — but not the shaded — metagenomes and became as abundant as — if not more than — the core genes (i.e. found in >90% of the genomes) of the *Salinibacter* population (Figure 3). Notably, function prediction analysis revealed that the majority of these genes are likely involved in sensing the environment, metabolite transport in/out of the cell, and gene regulation such as various glycosyltransferases [13••,26••], although the exact functional role of most of these genes remains elusive at present. Further, we have previously observed a correlation between the glycosyltransferase genome load and the fitness in salinities below saturation [29•]. Collectively, these results indicate that while the majority of the (non-core) pangenome may be ephemeral and/or functionally neutral under the prevailing conditions in this study, at least a small fraction of it appears to be important when the population adapts to environmental transitions, and we have outlined an approach to detect such genes (Figures 1 and 3). It is important to note that additional genes could be ecologically important during different environmental transitions such as viral predation and fall nutrient flux due to rain events, and these genes remain to be identified.

Remarkably, similar gene abundance patterns during the same treatments in salinity and light intensity were observed with the major archaeal sequence-discrete population of the salterns, *Haloquadratum walsbyi* [26••], revealing commonalities in how the two domains of life cope with changing conditions. Notably, the results for *Hqr. walsbyi* were based on population genome binning due to the fastidious nature of this archaeon in lab cultivation [26••], suggesting that culture-independent approaches could also be used to identify the ecologically important pangenome genes when populations are abundant enough to be adequately sampled by sequencing (e.g. >5–6X coverage [10•]). In addition, the ecologically important genes were carried by distinct *Hqr. walsbyi* subpopulations, or ecotypes, in terms of their intra versus inter-population ANI.

Figure 3



Abundance dynamics of non-core pangenome genes observed during environmental transition.
Each circle represents a genome and shows the fraction of core (in red) and variable (strain-specific and rare gene; in blue) genes in the genome. Yellow denotes the fraction of the variable genes that are ecologically important for the environmental perturbation (transition) that the population undergoes (middle). Note that the relative abundance of the ecologically important genes becomes high during transition before returning to low, post-transition. This pattern summarizes the dynamics observed for a few specific genes in extreme halophilic bacterial and archaeal species when they underwent a salinity change (representative functions are provided in the Figure key). In the *Hqr. walsbyi* case, the ecologically important genes are carried by distinct subpopulations, which was not observed in *Sal. ruber* (see text for further details).

Hence, in the *Hqr. walsbyi* case, and unlike *Sal. ruber*, it appears that the adaptation to different salinity concentration has led to subpopulation differentiation and (ongoing) speciation. In *Sal. ruber* case, the strains that carry the ecologically important genes for the environmental transition appear to be highly similar, if not identical, in other parts of the genome and thus, it has been challenging to document strain or subpopulation replacement based on short-read data [13••]. Long-read data and additional experiments will be required to provide further insights in the *Sal. ruber* case with respect to strain replacement. It would also be interesting to understand the underlying mechanism(s) that have led to the two different outcomes in terms of intra-species diversity in *Hqr. walsbyi* versus *Sal. ruber* in the future.

Conclusions

Recent studies have shown that microbial communities are predominantly composed of sequence-discrete populations, defined usually — but not necessarily always — at 95% ANI level, which can serve as an important unit to study community function and adaptation. The methods outlined above, and in Figure 1, provide the means to identify, track and study these populations over time or perturbations. Our own results with *Sal. ruber* and *Hqr. walsbyi* show that such populations harbor extensive gene content diversity (open pangenome), the majority of which may be neutral or ephemeral but at least a small minority of the (non-core) genes appear to be important for ecosystem functioning and population adaptation to changing conditions. These results are also consistent with prevailing theories about the functional role of the pangenome diversity [1,16]. It would be interesting to test how these results with the *Sal. ruber* and *Hqr. walsbyi* genomes from the salterns apply, or not, to other, more complex habitats in the future.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors would particularly like to thank especially the whole team at Salinas d'Es Trenc and Gusto Mundial Balearides, S.L. (Flor de Sal d'Es Trenc) for allowing access to their facilities and for their support in performing the experiments; and to Bonnie Baxter and Brian Hedlund for providing the sample of the Great Salt Lake included in Figure 1. This work was partly funded by the US National Science Foundation, awards #1831582 and #1759831 (to KTK), and by the projects CLG2015_66686-C3-1-P, PGC2018-096956-B-C41 and RTC-2017-6405-1 of the Spanish Ministry of Science, Innovation and Universities (to RRM), which were also supported with European Regional Development Fund (FEDER) funds. RRM acknowledges the financial support of the sabbatical stay at Georgia Tech by the grant PRX18/00048 also from the Spanish Ministry of Science, Innovation and Universities.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Konstantinidis KT, Ramette A, Tiedje JM: **The bacterial species definition in the genomic era**. *Philos Trans R Soc Lond B Biol Sci* 2006, **361**:1929-1940.

2. Konstantinidis KT, Tiedje JM: **Genomic insights that advance the species definition for prokaryotes**. *Proc Natl Acad Sci U S A* 2005, **102**:2567-2572.

3. Doolittle WF: **Speciation without species: a final word**. *Philos Theor Pract Biol* 2019, **11**.

4. Rossello-Mora R, Amann R: **Past and future species definitions for Bacteria and Archaea**. *Syst Appl Microbiol* 2015, **38**:209-216.

5. Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL *et al.*: **Opinion: re-evaluating prokaryotic species**. *Nat Rev Microbiol* 2005, **3**:733-739.

6. Caro-Quintero A, Konstantinidis KT: **Bacterial species may exist, metagenomics reveal**. *Environ Microbiol* 2012, **14**:347-355.

7. Konstantinidis KT, DeLong EF: **Genomic patterns of recombination, clonal divergence and environment in marine microbial populations**. *ISME J* 2008, **2**:1052-1065.

8. Bendall ML, Stevens SL, Chan LK, Malfatti S, Schwientek P, Tremblay J, Schackwitz W, Martin J, Pati A, Bushnell B *et al.*: **Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations**. *ISME J* 2016, **10**:1589-1601.

9. Johnston ER, Rodriguez RL, Luo C, Yuan MM, Wu L, He Z, Schuur EA, Luo Y, Tiedje JM, Zhou J *et al.*: **Metagenomics reveals pervasive bacterial populations and reduced community diversity across the Alaska tundra ecosystem**. *Front Microbiol* 2016, **7**:579.

10. • Meziti A, Tsementzi D, Rodriguez RL, Hatt JK, Karayanni H, Kormas KA, Konstantinidis KT: **Quantifying the changes in genetic diversity within sequence-discrete bacterial populations across a spatial and temporal riverine gradient**. *ISME J* 2019, **13**:767-779

Addressed several methodological limitations of metagenome-assembled genomes (MAGs) based on comparisons of MAG and isolate genomes originating from the same sample, and provided guidelines on how to detect and deal with these issues.

11. Luo C, Walk ST, Gordon DM, Feldgarden M, Tiedje JM, Konstantinidis KT: **Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species**. *Proc Natl Acad Sci U S A* 2011, **108**:7200-7205.

12. • Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S: **High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries**. *Nat Commun* 2018, **9**:5114

Showed that the ANI gap between species is robust by comparing all prokaryotic genomes publicly available, and provided a new kmer-based ANI implementation (FastANI) to allow the efficient calculation of ANI values among thousands of genomes.

13. •• Conrad R, Viver T, Gago JT, Hatt JK, Venter F, Rosselló-Móra R, Konstantinidis KT: **Toward quantifying the adaptive role of bacterial pangenomes during environmental perturbations**. *bioRxiv* 2021

Quantified the gene content diversity within a natural bacterial population based on 100 isolate genomes of the population, and provided hypotheses and data on how immense intra-population gene diversity may be maintained in light of frequent environmental perturbations.

14. •• Viver T, Orellana L, Gonzalez-Torres P, Diaz S, Urdiain M, Farias ME, Benes V, Kaempfer P, Shahinpei A, Ali Amoozegar M *et al.*: **Genomic comparison between members of the Salinibacteraceae family, and description of a new species of Salinibacter (*Salinibacter altiplanensis* sp. nov.) isolated from high altitude hypersaline environments of the Argentinian Altiplano**. *Syst Appl Microbiol* 2018, **41**:198-212

This manuscript made the case that manmade salterns represent an ideal system to understand the mechanisms creating and maintaining microbial diversity, and provided a methodology to isolate in pure culture the major bacterial taxa of the saltern microbial communities.

15. •• Rodriguez RL, Jain C, Conrad RE, Aluru S, Konstantinidis KT: **Reply to: "re-evaluating the evidence for a universal genetic boundary among microbial species"**. *Nat Commun* 2021, **12**:4060

Showed that genomes of intermediate ANI relatedness (that is, pairs of genomes whose ANI values fall within the ANI gap observed between species at 85–95% ANI) are ecologically differentiated, on their way to speciation, and the low frequency of such genomes is due to (distinct) ecological preferences, rather than sampling biases. Hence, the ANI gap can be used to identify and track the populations (units) of microbial communities.

16. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS *et al.*: **Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome"**. *Proc Natl Acad Sci U S A* 2005, **102**:13950-13955.

17. Daubin V, Ochman H: **Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli***. *Genome Res* 2004, **14**:1036-1042.

18. Casamayor EO, Massana R, Benlloch S, Ovreas L, Diez B, Goddard VJ, Gasol JM, Joint I, Rodriguez-Valera F, Pedros-Alio C: **Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern**. *Environ Microbiol* 2002, **4**:338-348.

19. Gomariz M, Martinez-Garcia M, Santos F, Rodriguez F, Capella-Gutierrez S, Gabaldon T, Rossello-Mora R, Meseguer I, Anton J: **From community approaches to single-cell genomics: the discovery of ubiquitous hyperhalophilic Bacteroidetes generalists**. *ISME J* 2015, **9**:16-31.

20. Mora-Ruiz MDR, Cifuentes A, Font-Verdera F, Perez-Fernandez C, Farias ME, Gonzalez B, Orfila A, Rossello-Mora R: **Biogeographical patterns of bacterial and archaeal communities from distant hypersaline environments**. *Syst Appl Microbiol* 2018, **41**:139-150.

21. Anton J, Rossello-Mora R, Rodriguez-Valera F, Amann R: **Extremely halophilic bacteria in crystallizer ponds from solar salterns**. *Appl Environ Microbiol* 2000, **66**:3052-3057.

22. •• Viver T, Orellana LH, Diaz S, Urdiain M, Ramos-Barbero MD, Gonzalez-Pastor JE, Oren A, Hatt JK, Amann R, Anton J *et al.*: **Predominance of deterministic microbial community dynamics in salterns exposed to different light intensities**. *Environ Microbiol* 2019, **21**:4300-4315

This work is based on metagenomics to quantify how much deterministic (e.g. selection by the environment) versus stochastic (e.g. random birth/death, genetic drift) factors affect the microbial community composition shifts observed in solar saltern brines.

23. Oren A: **Life at high salt concentrations**. In *The Prokaryotes.* Edited by Rosenberg E, DeLong EF, Thompson F, Lory S, Stackebrandt E. Springer; 2013.

24. Kurth D, Amadio A, Ordonez OF, Albarracin VH, Gartner W, Farias ME: **Arsenic metabolism in high altitude modern stromatolites revealed by metagenomic analysis**. *Sci Rep* 2017, **7**:1024.

25. Anton J, Oren A, Benlloch S, Rodriguez-Valera F, Amann R, Rossello-Mora R: ***Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds**. *Int J Syst Evol Microbiol* 2002, **52**:485-491.

26. •• Viver T, Conrad RE, Orellana LH, Urdiain M, Gonzalez-Pastor JE, Hatt JK, Amann R, Anton J, Konstantinidis KT, Rossello-Mora R: **Distinct ecotypes within a natural haloarchaeal population enable adaptation to changing environmental conditions without causing population sweeps**. *ISME J* 2021, **15**:1178-1191

This manuscript shows, using a metagenomic approach, that distinct ecotypes, within the 95% ANI level, appear to exist for the dominant archaeal population in salterns, *Haloquadratum walsbyi*, and are adapted to distinct environmental conditions.

27. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R *et al.*: **The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates**. *J Bacteriol* 2008, **190**:6881-6893.

28. • Rodriguez-R LM, Tsementzi D, Luo C, Konstantinidis KT: **Iterative subtractive binning of freshwater chronoseries metagenomes identifies over 400 novel species and their ecologic preferences**. *Environ Microbiol* 2020, **22**:3394-3412

Provided novel bioinformatics approaches to determine presence/absence of a target genome in a metagenome, and quantify ecological preference and prevalence of a genome in time-metagenomic or spatial-metagenomic datasets.

29. • Peña A, Keeling H, Huertas-Cepas J, Santos F, Yarza P, Brito-Echeverría J, Lucio M, Schmitt-Kopplin P, Meseguer I,

Schenowitz C *et al.*: **Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains**. *ISME J* 2010, **4**:882-895
This manuscript showed at that time the closest genome to genome comparison between two *Salinibacter* strains isolated from the same sample. There we could observe how one strain, encoding for a genomic island with high number of glycosyltransferases exhibited a metabolome with higher amount of glycosilated residues in contrast to the closest relative type strain, and this correlated with a better performance in salinities below saturation of the former.