Genomic evidence for the degradation of terrestrial organic matter by pelagic Arctic Ocean Chloroflexi bacteria

David Colatriano1, Patricia Q. Tran1, Celine Guéguen2, William J. Williams3, Connie Lovejoy4,5 & David A. Walsh1

The Arctic Ocean currently receives a large supply of global river discharge and terrestrial dissolved organic matter. Moreover, an increase in freshwater runoff and riverine transport of organic matter to the Arctic Ocean is a predicted consequence of thawing permafrost and increased precipitation. The fate of the terrestrial humic-rich organic material and its impact on the marine carbon cycle are largely unknown. Here, a metagenomic survey of the Canada Basin in the Western Arctic Ocean showed that pelagic Chloroflexi from the Arctic Ocean are replete with aromatic compound degradation genes, acquired in part by lateral transfer from terrestrial bacteria. Our results imply marine Chloroflexi have the capacity to use terrestrial organic matter and that their role in the carbon cycle may increase with the changing hydrological cycle.
The Arctic Ocean accounts for 1.4% of global ocean volume but receives 11% of global river discharge. Up to 33% of the dissolved organic matter in the Arctic Ocean is of terrestrial origin and a major fraction of this terrestrial dissolved organic matter (tDOM) originates from carbon-rich soils and peatlands. With thawing permafrost and increased precipitation occurring across the Arctic, increases in freshwater runoff and riverine transport of organic matter to the Arctic Ocean are predicted, which will increase tDOM fluxes and loadings. The additional tDOM may represent new carbon and energy sources for the Arctic Ocean microbial community and contribute to increased respiration, which would result in the Arctic being a source of dissolved inorganic carbon to the ocean. Alternatively, it as moves from its source of origin to the Arctic Ocean tDOM could become more recalcitrant to bacterial metabolism and represent a long-term sequestration of the newly released carbon making the Arctic more carbon neutral. However, an estimated 50% of Arctic Ocean IDOM is removed before being released to the Atlantic, at least in part by microbial processes. As input of tDOM increases, knowledge on its microbial transformation will be critical for understanding changes in Arctic carbon cycling.

The marine SAR202 is a diverse and uncultivated clade of Chloroflexi bacteria that comprise roughly 10% of planktonic cells in the dark ocean. SAR202 is also common in marine sediments and deep lakes. It has long been speculated that SAR202 may have a role in the degradation of recalcitrant organic matter, and the recent analysis of SAR202 single-cell-amplified genomes (SAGs) lends support to this notion. More generally, Chloroflexi, including those in the SAR202 clade, are also present in the upper layers of the Arctic Ocean, leading to the hypothesis that recalcitrant organic compounds present in high Arctic tDOM could be utilized by this group.

**Results**

In this study, we analyzed Chloroflexi metagenome-assembled genomes (MAGs) generated from samples collected from the vertically stratified waters of the Canada Basin in the Western Arctic Ocean. A metagenomic co-assembly was generated from samples originating from the surface layer (5–7 m), the subsurface chlorophyll maximum (25–79 m) and a layer corresponding to the terrestrially-derived DOM fluorescence (FDOM) maximum previously described within the cold Canada Basin halocline comprised of Pacific-origin waters (177–213 m). The Pacific-origin FDOM maximum is due to sea ice formation and interactions with bottom sediments on the Beaufort and Chukchi shelves, which themselves are influenced by coastal erosion and river runoff. Binning based on tetranucleotide frequency and coverage resulted in 360 MAGs from a diversity of marine microbes. Six near-complete Chloroflexi MAGs were identified. Based on 16S rRNA gene phylogeny, these MAGs represented three distinct SAR202 subclades (SAR202-III-AAA240, SAR202-III-AAA240, SAR202-VII). The AncK29 clade and the TK10 clade were also present. The marine SAR202 MAGs were distinct from previously published MAGs from the Arctic Ocean populations. The SAR202-VII MAG was the only MAG that contained many genes implicated in the degradation of aromatic compounds.

**AROMATIC COMPOUNDS**

Aromatic compounds are abundant in marine systems and can be produced from a variety of sources, including terrestrial input and microbial degradation of recalcitrant organic compounds. The metabolic potential of SAR202 to utilize aromatic compounds was investigated by analyzing the presence of genes involved in the catabolism of aromatic compounds. The SAR202-VII MAG contained numerous genes implicated in the degradation of aromatic compounds, including enzymes involved in the metabolism of catechol, protocatechuic acid, and vanillyl acid. The MAG also contained genes for the catabolism of benzoate, salicylate, and naphthalene.

**AROMATIC RING-CLEAVING HYDROXYLASES**

A total of 42 ring-cleaving hydroxylases were annotated in the SAR202-VII MAG. These enzymes are involved in the degradation of aromatic compounds by hydroxylation at the aromatic ring. The SAR202-VII MAG contained numerous genes for aromatic ring-cleaving dioxygenases, including enzymes involved in the metabolism of chlorophenol, phenol, and naphthalene.

**RING-CLEAVING MONOOXYGENASES**

Aromatic compounds are metabolized by ring-cleaving monooxygenases, which can cleave the aromatic ring and release products that can be further degraded by other enzymes. The SAR202-VII MAG contained numerous genes for aromatic ring-cleaving monooxygenases, including enzymes involved in the metabolism of chlorophenol, phenol, and naphthalene.

**RING-CLEAVING DECARBOXYLASES**

Aromatic compounds are also metabolized by ring-cleaving decarboxylases, which can cleave the aromatic ring and release products that can be further degraded by other enzymes. The SAR202-VII MAG contained numerous genes for aromatic ring-cleaving decarboxylases, including enzymes involved in the metabolism of chlorophenol, phenol, and naphthalene.

**AROMATIC COMPOUND DEGRADATION IN SAR202-VII**

The SAR202-VII MAG contained numerous genes implicated in the degradation of aromatic compounds, including genes for the metabolism of catechol, protocatechuic acid, and vanillyl acid. The MAG also contained genes for the catabolism of benzoate, salicylate, and naphthalene. The presence of these genes suggests that SAR202-VII has the metabolic potential to access carbon and energy available in aromatic compounds in the Arctic Ocean.

The diversity of SAR202-VII MAGs implicated in aromatic compound degradation lead us to hypothesize that they may have
originated by lateral gene transfer from terrestrial bacteria. To test this, we targeted aromatic compound degradation genes (the ring-cleaving dioxygenases, specifically) in the Chloroflexi MAGs for in-depth phylogenetic analyses. The genomic diversity of marine Chloroflexi was expanded in our analysis by including 130 Chloroflexi MAGs recently assembled and binned from the TARA Oceans project\textsuperscript{29}. A number of the SAR202-VII-2 ring-cleaving dioxygenase homologs were most closely related to

*Fig. 1* Metagenomic survey of microbial diversity in the Canada Basin. **a** Sampling locations and environmental profiles of the Canada Basin. Sample locations and the associated environmental datasets were plotted using Ocean Data View version 4.7.7\textsuperscript{42}. **b** Concatenated protein phylogeny of 360 Arctic Ocean MAGs inferred by MetaWatt and visualized in iTOL. The three inner tracks present relative coverage of MAGs averaged across samples collected from surface waters, subsurface chlorophyll maximum (SCM) and the fluorescent dissolved organic matter maximum (FDOM max). The outer track presents estimated MAG completeness as inferred by MetaWatt. MAG completeness ranged from 25 to 94%
proteins from the TARA Ocean Chloroflexi MAGs and other marine originating genomes, particularly the catechol dioxygenases (Supplementary Figure 3), gentisate 1,2-dioxygenases (Fig. 4) and methylgallate dioxygenases (Supplementary Figure 4) indicating that aromatic compound degradation in Chloroflexi is not restricted to the Arctic Ocean. However, lateral gene transfer from terrestrial bacteria was also evident. For example, an annotated gentisate dioxygenase gene was positioned within a clade of terrestrial Actinomycetes (Fig. 4). Additional genes involved in the degradation of structures related to catechol, protocatechuate and gentisate were phylogenetically associated with homologs from terrestrial Acidobacteria (Supplementary Figure 5), Actinobacteria (Supplementary Figure 6), Armatimodades (Fig. 4), Delta-proteobacteria (Supplementary Figures 3

**Fig. 2** Diversity and distribution of Arctic Ocean Chloroflexi MAGs. a Maximum likelihood tree of Chloroflexi based on partial 16S rRNA gene sequences. Blue taxa are from Canada Basin metagenomes and red taxa are from the six Chloroflexi MAGs. Bootstrap values of >70% are shown (100 replicates). b Distribution of MAGs in the Canada Basin based on metagenome coverage at the surface, subsurface chlorophyll maximum (SCM) and fluorescent dissolved organic matter maximum (FDOM max). c Distribution of MAGs in global ocean metagenomes based on fragment recruitment. Two deep ocean SAR202 SAGs from Landry et al. 18 were included for comparison. Location of TARA ocean metagenomes (red) and bathypelagic metagenomes (green) are shown on the map generated using Ocean Data View version 4.7. 

18
Table 1 Genomic characteristics of MAGs

| MAG            | Size (Mb) | Cov (x) | GC (%) | Completeness (%) | Contamination (%) | N50 (kb) | # of Contigs |
|----------------|-----------|---------|--------|------------------|-------------------|----------|--------------|
| SAR202-II-3    | 1.36      | 68      | 39     | 80               | 0                 | 35       | 46           |
| SAR202-II-177A | 1.62      | 101     | 42     | 82               | 1                 | 36       | 52           |
| SAR202-VII-2   | 2.78      | 16      | 59     | 99               | 0                 | 246      | 8            |
| SAR202-VI-29A  | 1.52      | 16      | 46     | 97               | 2                 | 59       | 2            |
| TK10-74A       | 2.45      | 12      | 69     | 81               | 2.3               | 38       | 76           |
| Anck29-46      | 1.11      | 11      | 32     | 77               | 0                 | 75       | 22           |

Fig. 3 Aromatic compound degradation genes and pathways in Arctic Ocean Chloroflexi MAGs. 

- **a** Abundance of aromatic compound degradation genes in Chloroflexi MAGs from the Canada Basin (pie chart) and a breakdown of those found specifically in SAR202-VII-2 (column plot). 
- **b** Predicted aromatic ring-opening enzymatic reactions identified in SAR202-VII-2 with gene loci displayed in the colored boxes. Genes in green and blue boxes were most closely related to homologs from terrestrial or marine bacteria, respectively. Genes in the blue/green boxes were in clades containing diverse environmental bacteria. 
- **c** Examples of predicted aromatic ring-modifying enzymatic reactions identified in SAR202-VII-2.
Fig. 4 Maximum likelihood tree of predicted gentisate 1,2-dioxygenases. Bootstrap values of >60% are shown (100 replicates). Homologs from SAR202-VII-2 are highlighted in red and homologs from TARA Ocean MAGs are highlighted in blue.
and 6), Beta-proteobacteria (Supplementary Figure 7) and a clade of diverse terrestrial phyla (Supplementary Figure 8). Additionally, 2 genitase 1,2-dioxygenase genes and 1 protocatechuate dioxygenase ligB gene were phylogenetically associated to a clade of genes from both terrestrial Delta-proteobacteria and marine microbes (Fig. 4 and Supplementary Figure 8). These putative gene acquisitions were unlikely due to contaminating scaffolds because the genes were located on long scaffolds that were assigned to Chloroflexi with high confidence based on tetranucleotide frequencies and the phylogenetic identity of housekeeping genes. Such a phylogenetic pattern supports the hypothesis that marine Chloroflexi acquired the capacity for aromatic compound degradation, at least in part, by lateral gene transfer from terrestrial bacteria.

Conclusions

In total, these results are consistent with Chloroflexi having a role in tDOM transformation in waters of the Arctic Ocean. This is the first study to our knowledge to associate a specific microbial group with tDOM metabolism in the Arctic Ocean and it expands on recent studies contributing to our understanding of the metabolic diversity of the abundant yet uncultivated marine Chloroflexi. Moreover, lateral gene transfer from terrestrial bacteria appears to have contributed to the evolution of aromatic compound degradation capabilities within marine Chloroflexi, particularly in regions of the Arctic Ocean impacted by terrestrial input.

The majority of MAGs were restricted to the humic-rich Pacific-origin halocline of the Canada Basin, however it is the surface waters that will be most immediately affected by increased freshwater input. Hence, our initial observations suggest a need for further research on the distribution of tDOM-utilizing microbes in other Arctic water masses with an aim to establish how common and phylogenetically widespread tDOM metabolism is in the Arctic Ocean. These water masses could include coastal surface waters at the mouth of the Mackenzie River, as well as regions of differing DOM composition such as the East Greenland Current and regions of differing DOM composition such as the East Greenland Current and the Norwegian Sea.

Methods

Sampling and DNA extraction. Twelve samples for metagenomics were collected in September 2015 during the Joint Ocean Ice Study cruise to the Canada Basin. For each sample, 4–8 L of seawater was sequentially filtered through a 50 μm pore mesh, followed by a 3 μm pore size polycarbonate filter and a 0.22 μm pore size Sterivex filter (Durapore; Millipore, Billerica, MA, USA). Filters were preserved in RNAlater and stored at −80 °C until processed in the laboratory. DNA was extracted from the Sterivex filter using the following method: filters were thawed on ice and RNAlater was removed. The Sterivex was then rinsed twice with a sucrose-based lysis buffer, and filled with 1.8 mL of the lysis buffer. Filters were treated with 100 μL of 125 mM MgCl2−1 l, 150 μM of 10 μg mL−1 RNase A and left to rotate at 37 °C for 1 h. After incubation, 100 μL of 100 mM −1 proteinase K and 100 μL of 20% SDS was added. Filters were left to rotate for 2 h at 55 °C. Lysate was removed from the filters. Protein was precipitated and removed with 0.583 volumes of MPC Protein Precipitation Reagent (Epiconent, Madison, WI, USA) and centrifugation at 10,000×g for 4 °C for 10 min. The supernatant was transferred to a clean tube. DNA was precipitated with cold isopropanol, and resuspended in low TE buffer.

Metagenomic sequencing, assembly, annotation, and binning. DNA sequencing of 12 samples was performed at the Department of Energy Joint Genome Institute (Walnut Creek, CA, USA) on the HiSeq 2500–1TB (illumina) platform. Paired-end sequences of 2 × 150 bp were generated for all libraries. A metagenome co-assembly of all raw reads was generated using MEGAHIT with kmer sizes of 22, 31, 41, 51, and 61 with a memory limit of 220,103,123. Gene prediction from the resulting assemblies was performed using the DOE Joint Genome Institute’s Integrated Microbial Genomes (IMG) database tool (version 4.11.0)11. Metagenomic binning was performed on scaffolds ≥10 kb in length using MetaWATT. Relative weight of coverage binning was set to 0.75 and the optimize bins and polish bins options were set to on. The taxonomic identity of predicted genes was assessed using a concatenated phylogenetic tree based on 95 single copy conserved genes as implemented in MetaWATT. Visualization of the tree and mapping of data on to taxa was performed with iTOL33. Estimation of MAG completeness and contamination was performed using CheckM34. Six Chloroflexi MAGs were identified for further analysis based on the presence of a 16S rRNA gene, high completeness and low contamination. Manual curation of the six Chloroflexi MAGs was performed and suspected contaminating scaffolds (single copy genes most similar to non-Chloroflexi taxa) were removed prior to further analysis of MAGs.

16S rRNA phylogenetic analysis. Chloroflexi diversity in the metagenomic assembly was assessed by 16S rRNA gene analysis. All 16S rRNA genes in the co-assembly were assigned to taxonomic groups using mothur35 and the Wang method with a bootstrap value cutoff of 60%36. Chloroflexi 16S rRNA genes greater than 360 bp were included in a phylogenetic analysis with Chloroflexi reference sequences. A multiple sequence alignment was generated using MUSCLE (implemented in MEGA6)37. Phylogenetic reconstructions were conducted by maximum likelihood using MEGA6-v.6.0 and the following settings: general time reversible model, gamma distribution model for the rate variation with four discrete gamma categories, and the nearest-neighbor interchange (NNI) heuristic search method38 with a bootstrap analysis using 100 replicates.

Single protein and concatenated protein phylogenies. A concatenated protein phylogeny was constructed using 30 Chloroflexi reference genomes and the 6 Canada Basin Chloroflexi MAGs. Orthologous genes in the 36 Chloroflexi genomes were identified using ProteinOrtho38. Fifty orthologs present in at least 34 of the 36 genomes were selected for concatenated phylogenetic analysis (Supplementary Data 3). Each orthologous protein family was aligned using MUSCLE (implemented in MEGA6) and alignment positions were masked using the probabilistic masker ZORRO39, masking columns with weights <0.5. The concatenated alignment consisted of 14,815 amino acid positions. Phylogenetic reconstructions were conducted by maximum likelihood using MEGA6-v.6.0 and the following settings: JTT substitution model, gamma distribution with invariant sites model for the rate variation with four discrete gamma categories, and the nearest-neighbor interchange (NNI) heuristic search method38 with a bootstrap analysis using 100 replicates.

For phylogenetic analysis of ring-cleaving dioxygenase sequences identified in SAR202-VII-2, query sequences were searched against UniRef90 and 130 Chloroflexi MAGs constructed from the TARA Oceans dataset37. The TARA Oceans Chloroflexi MAGs were used as is with no manual curation. UniRef90 sequences and the top TARA Ocean MAG hits for each dioxygenase were aligned with their respective SAR202-VII-2 homologs with MUSCLE (implemented in MEGA6) and alignment positions were masked using the probabilistic masker ZORRO39, masking columns with weights less than 0.5. Phylogenetic reconstruction was conducted using the same settings as the concatenated phylogeny.

Comparative genomics and metabolic reconstruction. The distribution of orthologs across Arctic Ocean genomes, as well as the identification of orthologs shared with the deep ocean SAGs, was determined using proteinortho38. Inference of protein function and metabolic reconstruction was based on the IMG annotations provided by the JGI, including KEGG, Pfam, EC numbers, and Metacyc annotations. Metabolic reconstruction was also facilitated by generated pathway genome databases for each MAG using the pathologic software available through Pathway Tools41.

Metagenomic fragment recruitment. The distribution of the Canada Basin MAGs in the global ocean was determined using best-hit reciprocal blast analysis similar to Landry et al.18 Unassembled metagenomic data from 25 samples (Supplementary Data 1) was first recruited to the six Canada Basin Chloroflexi MAGs as well as two SAR202 SAGs originating from the deep North Pacific Ocean from the Hawaiian ocean time series (HOTS) and the deep North Atlantic Ocean15. Metagenomes from the TARA Ocean project used here were representative of the surface, chlorophyll maximum and mesopelagic waters from the North Atlantic, South Atlantic, North Pacific, Coastal North Pacific, South Pacific, Brazil and the Antarctic Peninsula. To reduce computational demand, only part 1 (1 Gbp of a random subset of reads) of each metagenomic dataset available at EBI was used (Supplementary Data 1). Additional bathypelagic metagenomes from the North Pacific and South Atlantic Oceans (LineP P04, P12 and P26, and Knorr S15 2500 m) were also included. All hits from the initial blast were then reciprocally queried against the Canada Basin Chloroflexi MAGs, bathypelagic SAR202 SAGs, and 130 Chloroflexi MAGs constructed from the
TARA oceans data. The best-hit was reported. Only hits with an alignment length ≥100 bp and a percent identity of 95% or more were counted (lower % identity cutoff did not alter the number of reads recruited in any significant manner). To compare the results among the different datasets, the number of recruited reads was normalized to total number of reads in each sample. The final coverage results were expressed as the number of reads per kilobase of the MAG per gigabase of metagenome (rpkgs).

Data availability

The metagenomic data generated in this study are available in the Integrated Microbial Genomes database at the Joint Genome Institute at https://img.jgi.doe.gov, GOLD Project ID: Ga0133547. Metagenome-assembled genome projects have been deposited at DDBJ/ENA/GenBank under the Bioproject PRNA471535 and accession numbers QGNM00000000 (for SAR202-II-3), QGNQ00000000 (for SAR202-II-177A), QGNQ00000000 (for SAR202-29A), QEVV00000000 (for SAR202-VII-2), QGPN00000000 (for Anck29-46), and QGNQ00000000 (for TK10-74A). The versions described in this paper are versions QGNM10000000 (for SAR202-II-3), QGNQ10000000 (for SAR202-II-177A), QGNQ01000000 (for SAR202-VI-29A), QEVQ01000000 (for SAR202-VII-2), QGPQ01000000 (for Anck29-46), and QGQ10100000 (for TK10-74A).

Received: 6 November 2017 Accepted: 24 May 2018
Published online: 05 July 2018

References

1. Carmack, E. C. et al. Freshwater and its role in the Arctic marine system: sources, disposition, storage, export, and physical and biogeochemical consequences in the Arctic and global oceans. J. Geophys. Res. Biogeosci. 121, 675–717 (2016).
2. Benner, R., Benitez-Nelson, B., Kaiser, K. & Amon, R. M. W. Export of young terrigenous dissolved organic carbon from rivers to the Arctic Ocean. Geophys. Res. Lett. 31, 1–4 (2004).
3. Opsahl, B., Benner, R. & Amon, R. M. W. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. Linnol. Oceanogr. 44, 2017–2022 (1999).
4. Bintanja, R. & Andry, O. Towards a rain-dominated Arctic. Nat. Clim. Change 7, 263–267 (2017).
5. Frey, K. E. & McColland, J. W. Impacts of permafrost degradation on arctic river biogeochemistry. HydroL Process 23, 169–182 (2009).
6. Vonk, J. E. et al. Activation of old carbon by erosion of coastal and subsea permafrost in Arctic Siberia. Nature 489, 137–140 (2012).
7. Jiao, N. et al. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat. Rev. Microbiol. 8, 593–599 (2010).
8. Hansell, D. A. & Carlson, C. A. Biochemistry of Marine Dissolved Organic Matter 2nd ed (Academic Press, Amsterdam, 2015).
9. Kaiser, K., Benner, R. & Amon, R. M. W. Fate of terrigenous dissolved organic carbon on the Eurasian shelves and export to the North Atlantic. J. Geophys. Res. 122, 4–20 (2017).
10. Giovannoni, S. J., Rappé, M. S., Vergin, K. L. & Adair, N. L. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria. Proc. Natl Acad. Sci. USA 93, 7797–7794 (1996).
11. Delong, E. F. et al. Community genomics among stratified microbial assemblages in the ocean's interior. Science 311, 496–503 (2006).
12. Morgan, R. M., Rappé, M. S., Urban, E., Conon, S. A. & Giovannoni, S. J. Prevalence of the chloroflexi-related SAR202 bacterioplankton cluster throughout the mesopelagic and deep sea. Appl. Environ. Microbiol. 70, 2836–2842 (2004).
13. Schrenk, M. et al. Latitude distribution of prokaryotic picoplankton populations in the Atlantic Ocean. Environ. Microbiol. 11, 2078–2093 (2009).
14. Varela, M. M., van Aken, H. M. & Herndl, G. J. Abundance and activity of Chloroflexi-type SAR202 bacterioplankton in the meso- and bathypelagic waters of the (sub)tropical Atlantic. Environ. Microbiol. 10, 1903–1911 (2008).
15. Urban, et al. Unusual bacterioplankton community structure in ultra-oligotrophic Crater Lake. Limnol. Oceanogr. 46, 557–572 (2001).
16. Urban, B., Vergin, K. L., Larson, G. L. & Giovannoni, S. J. Bacterioplankton communities of Crater Lake, OR: dynamic changes with eutrophic zone food web structure and stable deep water populations. Hydrobiologia 574, 161–177 (2007).
17. Yamasaki, T. & Sekiguchi, Y. Cultivation of uncultured Chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi “subphylum I” with natural and biotechnological relevance. Microbes Environ. 24, 205–216 (2009).
Authors contributions

D.C. and D.W. designed and carried out the metagenomic experiments. W.J.W., C.L. and C.G. designed the sampling strategy. W.J.W. provided oceanographic data. D.W. collected samples. D.C. extracted the environmental genomic DNA and analyzed sequencing data. P.T. contributed to the bioinformatic analyses. D.C. and D.W. with contributions from C.G. and C.L. wrote the manuscript. All authors commented on the manuscript.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-018-0086-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018