







## Environmental DNA signatures distinguish between tsunami and storm deposition in overwash sand

Wenshu Yap <sup>1,2</sup>, Adam D. Switzer <sup>1,2</sup>✉, Chris Gouramanis<sup>3</sup>, Ezequiel Marzinelli <sup>4,5</sup>, Winona Wijaya<sup>1</sup>, Yu Ting Yan<sup>1</sup>, Dale Dominey-Howes <sup>6</sup>, Maurizio Labbate <sup>7</sup>, Seshachalam Srinivasalu<sup>8</sup>, Kruawun Jankaew<sup>9,10,11</sup> & Federico M. Lauro <sup>1,5</sup>✉

Sandy onshore deposits from tsunamis are difficult to distinguish from storm deposits, which makes it difficult to assess coastal hazards from the geological record. Here we analyse environmental DNA from microbial communities preserved in known tsunami and storm-deposited sediments and intercalating soils and non-marine sediments near Cuddalore, India, and Phra Thong Island, Thailand. Both sites were impacted by the 2004 Indian Ocean Tsunami and a subsequent storm flooding event (2011 Cyclone Thane at Cuddalore and a 2007 storm at Phra Thong Island). We show that the microbial communities in the overwash deposits are significantly different from soil and sediments that are not derived by overwash processes at both locations. Our method also successfully discriminates between modern tsunami deposits and storm deposits. We suggest molecular techniques have the potential to accurately discriminate overwash deposits from catastrophic natural events.

<sup>1</sup>Asian School of the Environment, Nanyang Technological University, Singapore, Singapore. <sup>2</sup>Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore. <sup>3</sup>Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia. <sup>4</sup>School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia. <sup>5</sup>Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore. <sup>6</sup>Asia-Pacific Natural Hazards and Disaster Risk Research Group, School of Geosciences, The University of Sydney, Sydney, NSW, Australia. <sup>7</sup>School of Life Sciences, University of Technology Sydney, Sydney, NSW, Australia. <sup>8</sup>Institute for Ocean Management, Anna University, Chennai, India. <sup>9</sup>Department of Geology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand. <sup>10</sup>PTTEP, 555/1 Energy Complex, Vibhavadi-Rangsit Road, Chatuchak, Bangkok, Thailand. <sup>11</sup>Present address: PTTEP, 555/1 Energy Complex, Vibhavadi-Rangsit Road, Chatuchak, Bangkok, Thailand. ✉email: [aswitzer@ntu.edu.sg](mailto:aswitzer@ntu.edu.sg); [flauro@ntu.edu.sg](mailto:flauro@ntu.edu.sg)

The rapid growth of coastal populations significantly increases the exposure of people and assets to coastal hazards such as tsunamis and storm surges<sup>1,2</sup>. Therefore, assessing coastal hazard risks lies at the core of policy planning and sustainable coastal development<sup>3</sup>. One key piece of information to assess flooding risks in coastal areas is to examine the frequency and magnitude of past events<sup>3,4</sup>. The magnitude and impact of modern events are well captured by instrumental and observational data<sup>3</sup>. However, the information that these technological advances provide is insufficient as coastal hazard scientists attempt to extend the knowledge of hazard intensities into the historical past to derive recurrence rates of large, infrequent catastrophic events<sup>3,5</sup>. To address the inherent limitation that the insufficient timespan of observational data provides, the occurrence of extreme wave events is commonly inferred from sedimentary deposits in the geological records<sup>5–9</sup>.

**Tsunami and storm hazards in the geological record.** Significant debate exists within the geoscience community about how to differentiate between the sediment deposits left behind by tsunamis and storms<sup>10–13</sup>. Tsunami and storm surge deposits are frequently indistinguishable due to similarities in sedimentary characteristics, even though they are generated by different mechanisms<sup>12,14,15</sup>. For this reason, multiproxy approaches have been developed by combining different diagnostic features of tsunami and storm deposits. However, each technique has inherent limitations. For example, grain size analysis can be inconclusive<sup>9,16</sup>; microfossils such as diatoms, foraminifera, and ostracods may only be preserved for a relatively short time in tropical settings as compared to temperate locations<sup>17–19</sup>; and chemical elemental signals present in the geological record might be misleading if they have been modified or removed by natural processes such as precipitation<sup>20</sup> and microbial activities<sup>21</sup>, or if the elemental sources are ambiguous<sup>16</sup>. In addition, compiling multiple proxies might not unequivocally show that deposits were laid down by overwash processes<sup>22</sup>.

**The potential of microbial molecular techniques.** There is a considerable amount of literature describing how microbial communities in different environmental settings respond to abrupt and intense “pulse” disturbances (i.e. relatively short-duration and discrete environmental alterations)<sup>23</sup> due to catastrophic events such as coastal flooding<sup>21,24–29</sup>, wildfire<sup>30,31</sup>, earthquake<sup>32,33</sup>. The primary aim of such studies is commonly to investigate how microbes respond to a disturbance in nutrients<sup>21,28–30</sup>, water resources<sup>24,25</sup>, substrate availability<sup>25,27–32</sup>, or other physical environment changes<sup>23,31–33</sup>.

To date, several studies have focused on microbial community changes after a tsunami flooding event (Table 1). A number of studies utilize conventional culturing techniques that target the culturable fraction of the microbial communities<sup>24–27</sup>. However, the culturable fraction of the community is unlikely to contain a detailed signature of the type of overwash since only a few microorganisms (<5%) within natural microbial communities are culturable<sup>34,35</sup>. Somboonna et al.<sup>29</sup>, and Asano et al.<sup>21,28</sup> were pioneers in using metabarcoding to study tsunami deposits. They reported that microbial communities in the sediment samples on the coastal zone vary before and after tsunami events, implying that tsunami deposits may contain a distinctive microbial signature. While all the studies listed in Table 1 focused on either comparing between tsunami deposits and unflooded terrestrial sediment samples, or marine and intertidal sediment before and after the flooding event (except<sup>36,37</sup>), none investigated the microbial community changes within the geological record at the same location. Moreover, there are no studies that

have investigated the differences in the microbial community between tsunami and storm deposit.

In this study, we applied molecular techniques to examine the microbial signatures of pulsed overwash deposits in two locations and attempt to answer some key questions in coastal hazard studies: (i) Do microbes differ in respective locations and environments? If yes, can we identify individual tsunami and storm deposits from intercalating soil and sediments? (ii) If we can differentiate overwash deposits from intercalating soil and sediment, can we distinguish between modern tsunami and storm deposits? (iii) Is it possible to identify the source of overwash deposits using molecular techniques? (iv) Can we detect a global microbial indicator for overwash deposits?

To answer all these questions, we collected samples from one site each in India and Thailand (Fig. 1) that were both impacted by the 2004 Indian Ocean Tsunami (2004 IOT). At the Thailand site, the 2004 IOT deposits occurred in a swale, and the storm overwash formed during an unnamed tropical depression in 2007, and was deposited behind the modern berm on a sandy beach ridge (Fig. 1c). At the Indian site, the 2004 IOT deposits were deposited behind a sandy beach dune, that was subsequently inundated by Tropical Cyclone Thane in December 2011 (Fig. 1f). We selected the Thailand site because the 2004 IOT deposits is clearly demarcated by a sandy layer between two organic mud layers<sup>5,9</sup>, and which differ from the sandy sediments coastward of the swale (Fig. 1c). The Thailand site serves as a good site to benchmark our microbial metabarcoding approach to investigate geological records. Next, we examined sediment samples from the India site, in which different units are sandy and display limited sedimentary differences (Fig. 1f). The Indian site allows us to explore our microbial metabarcoding approach to compare with conventional sedimentological and stratigraphic methods and data, such as sediment grain-size characteristics (Fig. 1), and organic geochemistry, such as C, N, and S concentrations (Supplementary Fig. 4). We applied a simple, next-generation high-throughput microbial metabarcoding approach (Fig. 2) that utilizes the genetic material extracted directly from modern environmental samples to characterize the microbial assemblages that are present (both alive and dead organisms) in the deposits<sup>36</sup>. Knowing “what is there” in the geological record allows the investigation of whether microbial diversity and abundance are underpinned by geological (e.g., deposition, weathering) or ecological (e.g., carbon cycle, nitrogen cycle) processes, and whether the microbial community differences can be related to the type of deposit or tied to the deposits source.

**Study sites description.** We collected sediment samples from two well-researched study sites: Phra Thong Island, Thailand<sup>5,9,16</sup> and Cuddalore, Tamil Nadu, India<sup>38–40</sup> (Fig. 1). Both sites were impacted by the 2004 Indian Ocean Tsunami (2004 IOT) and by different storm events; a minor tropical depression in 2007 at Phra Thong Island, and a Category 2 tropical cyclone—Cyclone Thane in 2011 at Cuddalore. Thus, these two sites provide a rare opportunity to directly compare: (i) known tsunami and storm events that have impacted the same coastline and (ii) the same tsunami event that impacted two different sites with markedly different coastal morphologies (Fig. 1, Supplementary Fig. 1). The Phra Thong site contains a ridge and swale setting with a variety of environmental and sedimentological features including sandy beach ridges and dunes and organic-rich, muddy swales (Fig. 1d, Supplementary Fig. 1)<sup>5</sup>. In contrast, the Cuddalore system has lower environmental diversity and is dominated by the sandy sediments characteristic of a coastal beach-dune and associated

**Table 1 Summary of all microbial studies on tsunami deposits.**

Events	Specific site location (Country, bay/island/prefecture)	Sample descriptions	Storage condition	Methodology for microbial community	Key findings	Ref
2004 Indian Ocean Tsunami	India, Chennai coast of Bay of Bengal.	Marine sediments and seawater were collected in Nov. 2004 (26 days before tsunami), Dec. 2004 (5-10 h after tsunami), Jan. 2005 (7 days after tsunami), Jan. 2005 (14 days after tsunami), Jan. 2005 (21 days after tsunami), and Jan. 2005 (28 days after tsunami).	NA	Cultured-based isolation of microbes, fungi and actinomycetes.	<ul style="list-style-type: none"> <li>• In pre-tsunami samples, culturable microbes were higher in concentration in the seawater samples than the sediments, while fungi and actinomycetes were recorded only in the sediments.</li> <li>• After the tsunami, all marine samples had a remarkably increased culturable microbial population that declined in subsequent days and became similar to the number before the tsunami.</li> </ul>	24
2004 Indian Ocean Tsunami	Andaman and Nicobar Island, South Andaman: Mithakhari, Crickadabad, Guptapara, New Manglutan, and Lohabarrak.	Sediment samples were collected from 5 low-lying coastal areas in May 2005. Samples were collected from surface (0-15 cm) and subsurface horizon (15-30 cm).	NA	Cultured-based isolation of microbes, fungi and actinomycetes.	<ul style="list-style-type: none"> <li>• Tsunami-affected land had a higher pH in the surface soil layer compared to subsurface as opposed to land not affected by tsunami.</li> <li>• Soil microbes in unaffected land had a higher culturable microbial population compared to tsunami-affected land.</li> <li>• Area that was once submerged by tsunami had a higher culturable microbial population in the soil after the water had permanently receded compared to areas that were permanently submerged.</li> </ul>	25
2004 Indian Ocean Tsunami	India, southeast coast between Vanagiri and Nagoor coastal sediments.	Intertidal marine sediments (0-15 cm) were collected from 12 sampling location pre- and post-tsunami event.	NA	Cultured-based isolation of microbes and fungi.	<ul style="list-style-type: none"> <li>• Pre-tsunami sediment samples had higher culturable microbial diversity than post-tsunami samples.</li> <li>• There were more pathogenic species (bacteria and fungi) in the pre-tsunami sediments.</li> </ul>	26
2004 Indian Ocean tsunami	Thailand, Phang Nga province, Phra Thong Island.	Sediment samples were collected in March 2011 from 0.40 km inland that was affected by the tsunami (S1) and 2.26 km inland that was not affected by the tsunami event (S2). S1 samples comprised 2004 Indian Ocean Tsunami (14.5 cm), 1-300 year old (22 cm), 300-600 year old (29 cm), 600 year old (38 cm) and >600 year old (46 cm). The number in parentheses represents the depth level where the sample was collected. S2 comprised samples collected from approximately the same depth as S1.	4 °C	Metabarcoding 16S rRNA <sup>b</sup> and 18S rRNA <sup>b</sup> using 454 pyrosequencing technique.	<ul style="list-style-type: none"> <li>• S1 comprised richer archaeal populations (2.36%) but lower diversity for fungi (1.30%) and protists (0.73%) as compared to S2 (archaea 0.58%; fungi 4.88%; protist 0.80%). The phyla and species distribution were different in S1 and S2.</li> <li>• Based on potential metabolic functions in the microbial communities using MetaGenomics-Rapid Annotation using Subsystems Technology (MG-RAST), the microbial community metabolic system in both locations were different. S1 was predominantly advanced metabolic subsystems of regulating and cell signaling, cell wall and capsule, protein metabolism, sulfur metabolism and carbohydrate whereas S2 carried high metabolic potentials for pathways of respiration, photosynthesis, and drug and bioactive compound production.</li> <li>• The habitat prediction based on percent of species indicators for marine, brackish, freshwater and terrestrial niches indicated that S1 largely consisted of marine-habitat indicator species.</li> </ul>	29
2011 Tohoku tsunami	Japan, Miyagi Prefecture, Sendai city, Hiyoriyama and Amamiya.	Sediment samples were collected in July 2012 from Hiyoriyama located 0.5 km inland from the coastline and was affected by the tsunami event, and Amamiyama located 12 km inland from the coastline and was not affected by the tsunami event. Water samples were collected 110 km and 150 km off the coast of Sendai city in August 2012, during the KT-12-21 cruise of R/V Tansei-Maru (JURCAOS, JAMSTEC).	-80 °C	Cultured-based bacterial isolation. Culture-independent Metagenomic sequencing using 454 pyrosequencing technique. <i>Arthobacter</i> isolate sequenced using 454/Ion PGM/PacBio sequencing technique.	<ul style="list-style-type: none"> <li>• Tsunami-affected soil has more strains isolated than the unaffected soil.</li> <li>• The isolated strains were found to belong to a single genus, <i>Arthrobacter</i>. This genus has a losses gene that responds to adaptation to an environment with high-iron concentration and the tsunami-affected soil were found to be rich in iron.</li> </ul> <p>Metagenomic data show that tsunami-affected soil samples have an over-representation of denitrification-related gene as well as the presence of pathogenic and marine bacterial genera and salt tolerant bacteria.</p>	27

**Table 1 (continued)**

Events	Specific site location (Country, bay/island/prefecture)	Sample descriptions	Storage condition	Methodology for microbial community	Key findings	Ref
2011 Tohoku Tsunami	Japan, Miyagi Prefecture, Higashimatsushima City rice field.	Sediment samples were collected in April 2012 from unflooded field, field flooded for 2 weeks and field flooded for 2 months.	4 °C	Metabarcoding 16S rRNA <sup>b</sup> using 454 pyrosequencing technique.	<ul style="list-style-type: none"> <li>Unflooded soil had higher microbial diversity compared to the flooded soil.</li> <li>Hierarchical clustering showed that the community structure of the soil bacteria in flooded soil (both 2 weeks and 2 months after) was clearly different from the unflooded soil.</li> <li>The effects of the tsunami on soil bacteria in agriculture fields may have lasted at least 1 year.</li> <li>Sulfur-oxidizing bacteria (SOB), nitrite-oxidizing bacteria (NOB) and zeta-proteobacteria may serve as an indicators of seawater inundation.</li> </ul>	28
2011 Tohoku tsunami	Japan, Hokkaido Island, Urahoro.	A 1.5 m-long sediment core was retrieved and sampled in the field. Seven samples from palaeotsunami sandy layers and from intercalating peat layers were collected. Three additional samples were taken from the surface sandy beach and a nearby natural wetland.	−20 °C	Micropalaeontological analyses of foraminifera and diatoms. Metabarcoding foraminifera-specific gene using Illumina sequencing technique.	<ul style="list-style-type: none"> <li>172 diatom species were identified with limited fossilized foraminifera tests and radiolaria detected from the samples.</li> <li>Foraminifera DNA<sup>a</sup> sequences were detected in all modern samples and most of the palaeo-deposits.</li> <li>The majority of the foraminifera DNA<sup>a</sup> belonged to brackish and marine taxa.</li> </ul>	36
2011 Tohoku tsunami	Japan, Miyagi Prefecture, Higashimatsushima City rice field.	Sediment samples were collected in April 2013, 2014 and 2015 from unflooded field, field flooded for 2 weeks, field flooded for 2 months and long-term flooded and cultivated field.	−30 °C	Metabarcoding 16S rRNA <sup>b</sup> using 454 pyrosequencing technique.	<ul style="list-style-type: none"> <li>There is limited microbial differences between field flooded for 2 months and unflooded field.</li> <li>There were higher number of sulfur-oxidizing bacteria (SOB) and halotolerant SOB in the field flooded for 2 months.</li> <li>Tsunami flooding created a unique environment that house halotolerant bacteria that are not found in marine sediments or agricultural soils.</li> </ul>	21

<sup>a</sup>DNA is deoxyribonucleic acid, here refer as genetic material extracted from the environmental samples that were later sequenced to generate the DNA sequences data  
<sup>b</sup>rRNA is ribosomal ribonucleic acid, it is commonly used in microbiology studies as a tool to characterize unknown microbial organisms

backshore system (Fig. 1g, Supplementary Fig. 1)<sup>38,39</sup>. Detailed geomorphological description is included in the methodology section.

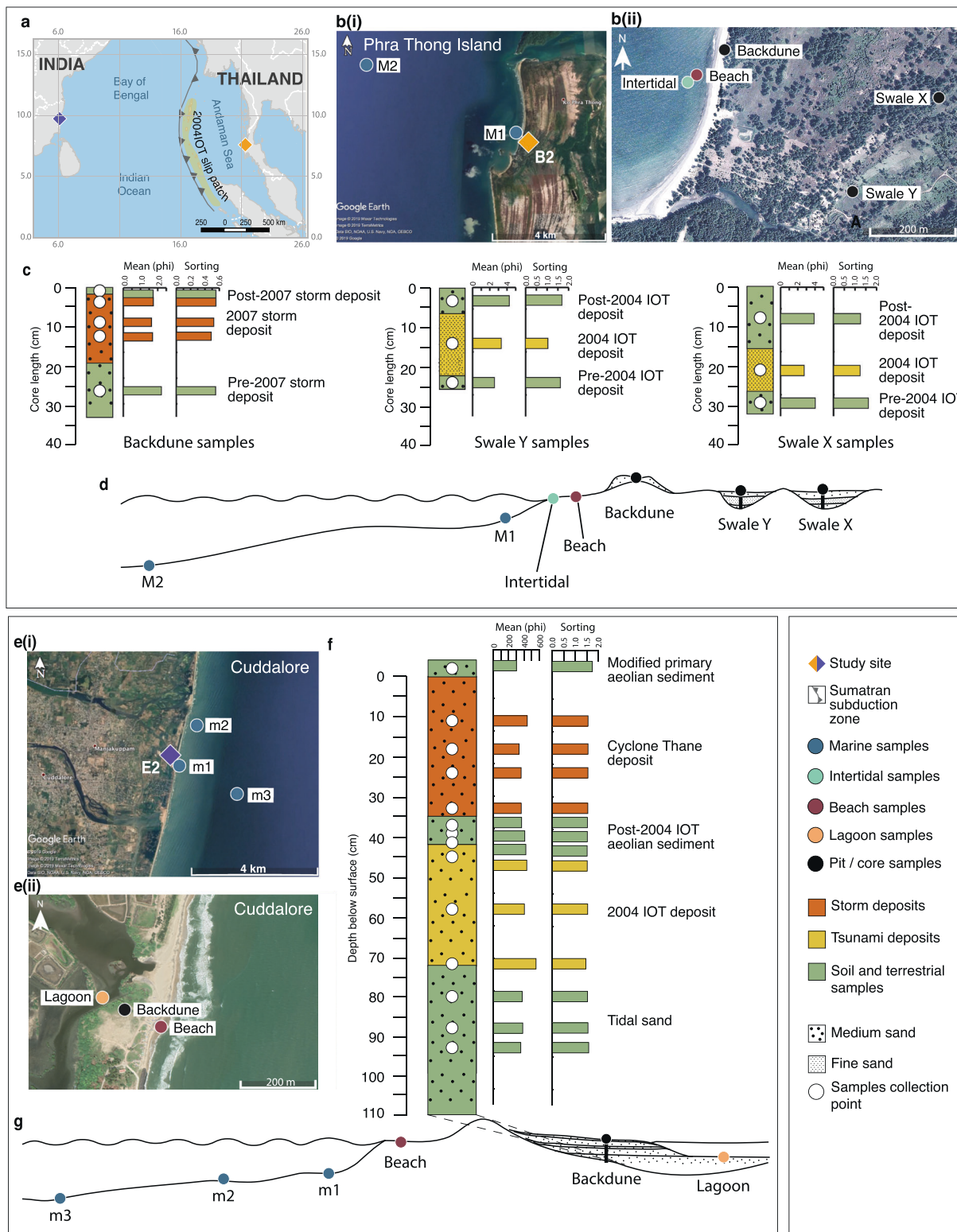
## Results

**Microbial community composition and diversity within samples.** We analyzed the microbial community present in two sediment cores, two beach sand, two intertidal sand, and two marine sediments (total 21 sediment samples) collected from Phra Thong Island and one sediment core, two beach sand, two lagoon sediments, and 6 marine sediments (total 36 sediment samples) and 16 water samples from Cuddalore (Supplementary Data 1). All samples were analyzed with metabarcoding of the SSU rRNA gene which is a universal genetic marker. The metabarcoding generated an average of 84,011 DNA sequences that were grouped into a total of 25,034 amplicon sequence variants (ASVs).

Microbial community diversity at each site was calculated after the dataset was rarefied to the minimum number of sequences 42,412 to simulate an even number of sequences per sample. We applied the Shannon ( $H'$ ) and Simpson's ( $D$ ) diversity indices to measure the abundances and evenness of the microbial community present within each sample<sup>41,42</sup>. Both diversity indices show that the overall microbial community diversity on Phra Thong Island ( $H' = 6.3283$ ,  $D = 0.9952$ ; Fig. 3) was higher than in Cuddalore ( $H' = 5.2393$ ,  $D = 0.9795$ ; Fig. 3). At both locations, on average, the microbial diversity was higher in the marine sediments (Phra Thong Island:  $H' = 6.6750$ ,  $D = 0.9974$ ; Cuddalore:  $H' = 6.1859$ ,  $D = 0.9961$ ; Fig. 3). Samples collected along the

coastline, i.e. intertidal sands ( $H' = 6.7502$ ;  $D = 0.9981$ ; Fig. 3) and backdune pit samples ( $H' = 6.4131$ – $6.5173$ ,  $D = 0.9925$ – $0.9967$ ; Fig. 3) on Phra Thong Island, and beach samples in Cuddalore ( $H' = 6.3438$ ,  $D = 0.9966$ ; Fig. 3) had higher microbial diversity than the organic-rich swale samples on Phra Thong Island ( $H' = 5.9358$ – $6.2960$ ,  $D = 0.9925$ – $0.9963$ ; Fig. 3). Overwash deposits at both locations had the lowest community richness (Phra Thong Island:  $H' = 6.2960$ – $6.4131$ ,  $D = 0.9939$ – $0.9967$ ; Cuddalore:  $H' = 4.8647$ – $5.1262$ ,  $D = 0.9727$ – $0.9777$ ; Fig. 3) when compared to the overlying and underlying units and other samples collected onshore. The microbial community preserved in the Cyclone Thane deposits have higher diversity ( $H' = 5.1262$ ;  $D = 0.9777$ ; Fig. 3) than the 2004 IOT deposits ( $H' = 4.8647$ ;  $D = 0.9727$ ; Fig. 3) at Cuddalore, whereas on Phra Thong Island, the microbial diversity for both tsunami and storm deposits were similar ( $H' = 6.2960$ – $6.4131$ ,  $D = 0.9939$ – $0.9967$ ; Fig. 3).

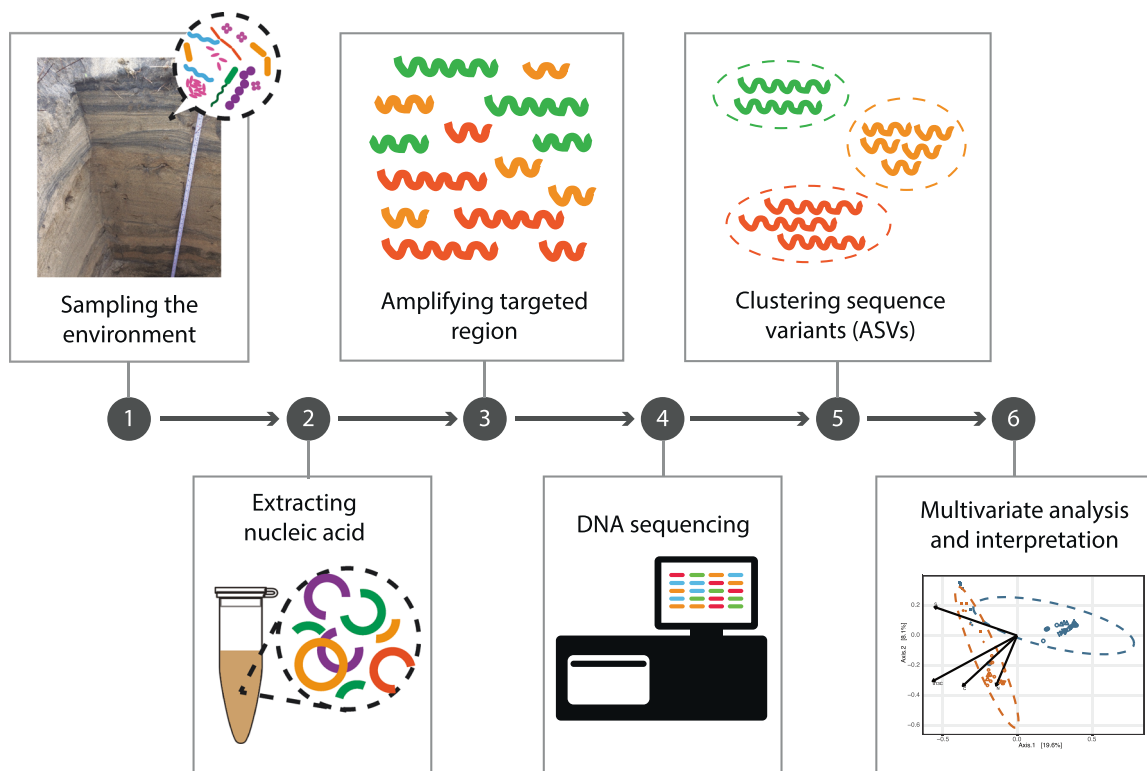
Overall, the bacterial communities as determined from 16S rRNA present in all the samples were predominantly of the phylum Proteobacteria (Supplementary Fig. 2). At Phra Thong Island, there was a large variety of Proteobacteria classes, including Gammaproteobacteria, Deltaproteobacteria, Alphaproteobacteria, and Betaproteobacteria (Supplementary Fig. 2). Conversely, and likely reflecting the lower diversity, the phylum Proteobacteria at Cuddalore were mainly from two classes: Gammaproteobacteria and Alphaproteobacteria (Supplementary Fig. 2). Within the kingdom Archaea, the phylum Thaumarchaeota were represented by order Nitrososphaerales and order Nitrosopumilales at both locations, with a much



lower proportion on Phra Thong Island (Supplementary Fig. 2). The eukaryotic community as determined from 18S rRNA in both locations was dominated by the phylum Ciliophora, class Colpodea and class Spirotrichea, and the phylum Cercozoa, class Filosa (Supplementary Fig. 2). As Phra Thong Island and Cuddalore harbored very different archaeal and bacterial communities, we adopt a consistent approach but analyzed the microbial community structure at each location independently.

**Microbial community dissimilarity between environments.** We applied principal coordinate analysis (PCoA) with a Bray–Curtis dissimilarity matrix to illustrate the difference between samples collected from (i) Phra Thong Island and Cuddalore and (ii) different environments at each site; marine sediments collected offshore, beach and intertidal sand collected onshore, lagoon sediment samples, and pit and core samples from backdunes and swales for both 16S and 18S rRNA data (Fig. 1, Supplementary Figs. 3, 6). For bacterial and archaea database (Supplementary

**Fig. 1** The study sites at Phra Thong Island (Thailand) and Cuddalore (India) that were both impacted by the 2004 Indian Ocean Tsunami (2004 IOT). Subsequently, in 2007, an unnamed tropical depression caused a storm overwash in the Thailand and in December 2011, the Indian site was inundated by a very severe tropical cyclone Thane. Location map ((a) is created in QGIS v3.10 while (b) and (e) are satellite images courtesy of Google Earth Pro v7.3.3.7786, Image data © 2019 Terra Metrics, Maxar Technologies) of Phra Thong Island (a, b(i), b(ii)) and Cuddalore, India (a, e(i), e(ii)). The stratigraphy (created in Adobe Illustrator v25.2.1) and the microbial sample collection depths (white dots) from a backdune pit and sediment cores from two swales (Swale X and Swale Y) are illustrated in (c) (top panel). f (bottom panel) shows the stratigraphy and microbial sample depths (white dots) from the backdune pit at Cuddalore. The grain size analysis (Mean (phi) and sorting) of the sediments are included on the right side of the respective stratigraphic diagrams. Environmental samples such as beach, intertidal, lagoon (only for Cuddalore), and marine sediment samples were also collected from both sites (Phra Thong Island: b(ii), d; Cuddalore: e(ii), g). Marine sediment and water samples were collected at 2 m (M1) and 20 m (M2) water depths offshore of Phra Thong Island, and marine sediment and water samples from offshore of Cuddalore were collected at 5 m (m1), 9.5 m (m2), and 15 m (m3) water depths. A schematic cross-section of the study site (created in Adobe Illustrator v25.2.1) is included on the bottom panel for Phra Thong Island (d) and Cuddalore (g). A detailed sample description is listed in Supplementary Data 1.

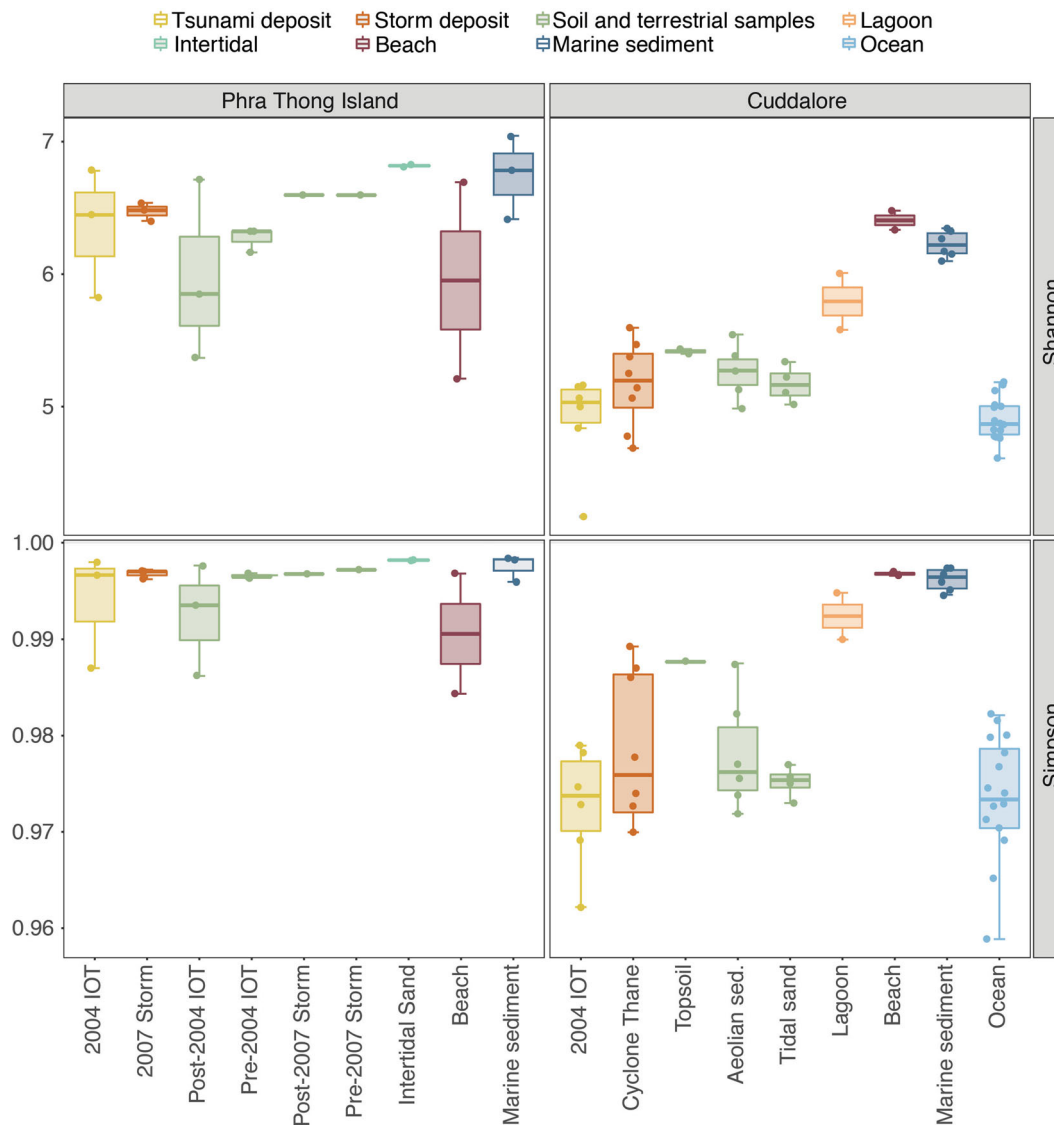


**Fig. 2** Schematic diagram of the environmental DNA metabarcoding workflow. Microbial metabarcoding approaches using genetic material extracted from environmental samples. A universal genetic marker (i.e. primer) was used to amplify a small subunit of ribosomal ribonucleic acid (SSU rRNA). SSU rRNA is stable and conserved for all organism genetic material that is functionally and evolutionarily similar, yet highly variable across different species. The amplified deoxyribonucleic acid (DNA) (also referred to as amplicons) were then sequenced through Next-Generation Sequencing (NGS) facility that generates high-throughput DNA sequence data. The data are processed through a bioinformatic pipeline that: (i) filters and trims low-quality sequences, (ii) corrects errors generated during sequencing (DADA2) and (iii) clusters the sequence variances (referred to as amplicon sequencing variants (ASVs)) based on the error model. These ASVs are used to infer the biological identity of members of the microbial communities. The differences between communities in the environmental samples and overwash sediments are then visualized using ordination techniques and tested statistically.

Fig. 3), the first PCoA axis in Supplementary Fig. 3a supports 20.9% of the variance and revealed a clear distinction between the microbial communities in marine water samples versus those in marine sediment samples. The second PCoA axis in Supplementary Fig. 3a (12.8% of the variance) shows a clear separation of the microbial communities based on location. In subsequent analyses, we excluded water samples and lagoon sediment samples from the Cuddalore dataset to standardize the testing variables for both locations. Subsequent PCoA with the smaller sample dataset (Supplementary Fig. 3b) shows that at both locations, the microbial communities from marine, beach and intertidal sediments (Supplementary Fig. 3b, filled shape) were dissimilar from samples collected from the backdunes and swales

(Supplementary Fig. 3b, outlined circle and triangles; Axis 2 = 8.1% of the variance). We observed that overwash deposits have a high similarity with the microbial communities in the overlying and underlying soil (Supplementary Fig. 3b).

Likewise, the PCoA calculated based on the 18S rRNA database (Supplementary Fig. 4) shows that the marine water samples were clustered separately from sediment samples (Supplementary Fig. 4a, Axis 1 = 15% of the variance), and the microbial dissimilarity between sediment samples was separated based on the sampling locations (Supplementary Fig. 4a, Axis 2 = 10.5% of the variance). As we remove water samples from the analysis and focuses on examining the microbial dissimilarity of the eukaryotic community in the sediment samples (Supplementary Fig. 4b), we



**Fig. 3** Boxplot of Shannon and Simpson's diversity indices with 95% confidence interval error bars, characterizing the richness and evenness of ASVs that were present on Phra Thong Island and at Cuddalore.

Shannon and Simpson diversity indices indicate that the richness and evenness of ASVs on Phra Thong Island were higher than at Cuddalore. Marine sediments have the highest diversity compared to other sample types (intertidal, beach, overwash deposits and intercalating soil and terrestrial samples) at both locations, followed by the beach and intertidal sand. Backdune samples from Cuddalore and swale samples from Phra Thong have relatively lower diversity than the marine and beach samples at both locations. On Phra Thong Island, the 2007 storm and 2004 IOT deposits have relatively higher diversity compared to the soils above and below the two event deposits. In contrast, Cyclone Thane and the 2004 IOT deposits at Cuddalore have the lowest diversity compared to other sediment samples collected at the same study location.

observed that the eukaryotic community in the Phra Thong Island samples were clustered together. Even though there is a faint separation between samples collected from the sandy backdunes (outlined shapes) and marine sediments and beach samples (filled shapes) in Cuddalore dataset, multivariate analysis on the 18S rRNA database shows no significant differences between marine sediments, beach, overwash deposits and intercalating soil and terrestrial samples (Supplementary Fig. 4).

**Effects of geochemistry and grain size on microbial dissimilarity?** To assess whether the dissimilarity between the microbial communities was affected by local geochemistry or nutrient availability in the sediments, we measured the three most important chemical elements in soil—specifically carbon, nitrogen, and sulfur<sup>43</sup> from each sediment sample. We performed linear regression analysis to determine the relationship between

each chemical variable with each ordination axes and projected this relationship onto the PCoA plot (Supplementary Fig. 3b). The results revealed that the chemical data were significantly correlated (Supplementary Table 1; *p*-value <0.02) to the dissimilarity between sampling sites but not between sample types at each site (Supplementary Fig. 3b).

On Phra Thong Island, soil samples extracted from the cores collected from muddy Swale X and Swale Y (further inland) had the highest total organic carbon (TOC) content (0.63–8.74%), whilst sandy backdune pit samples (nearer to coast where storm samples were collected) had a relatively lower TOC content (0.24–0.54%; Supplementary Fig. 5). A similar trend was observed for total nitrogen (TN) and total sulfur (TS) data on Phra Thong Island, where the core samples from Swale X and Swale Y had higher TN (0.01–0.75%) and TS (0.01–0.26%) compared to the backdune samples that had minimal amounts of TN ( $\leq 0.02\%$ ) and TS ( $\leq 0.01\%$ ) (Supplementary Fig. 5). In the sedimentary

layers and soils of the swales, the 2004 IOT deposits had lower TOC (<2%), TN (<0.04%) and TS (<0.05%) than the post-2004 IOT soil above (TOC: 0.63–8.74%; TN: 0.03–0.75%; TS: 0.01–0.26%) and the pre-2004 IOT soil below it (TOC: >3%; TN: 0.01–0.32%; TS: ≥0.08%) (Supplementary Fig. 5).

The Cuddalore system had a much simpler geochemical profile as TN or TS were not detected in the Cyclone Thane deposits, 2004 IOT deposits and the majority of the intercalating sand layers (Supplementary Fig. 5). The one exception was the post-2004 IOT aeolian deposit (refers to sediment that was transported and deposited by the wind) which had TN of approximately 0.01% (Supplementary Fig. 5). The TOC in the Cuddalore sediment samples was also consistently low (0.03–0.11%), with the exception of marine sediments that had a slightly higher amount of TOC ranging from 0.25% to 1.53% (Supplementary Fig. 5).

To investigate grain-size variation (Fig. 1) and the possible link to ASVs we ran an Envfit analysis with the inclusion of grain size distribution data (mean and sorting) shown in Supplementary Table 2. Supplementary Table 2 shows that there is a significant correlation ( $p$ -value <0.05) that links the environmental DNA to grain-size mean or sediment sorting. As we focus on understanding the microbial dissimilarity between each stratigraphic unit within the sediment profiles, i.e. examining the overwash deposits versus intercalating soil and terrestrial samples, we observed that there was no correlation between microbial dissimilarity and mean grain size (Supplementary Table 2).

**Distinguishing between terrestrial soil, tsunami, and storm deposits.** In order to resolve the complex relationship between storm and the 2004 IOT deposits, and intercalating soil and sand layers, we performed a constrained ordination using distance-based redundancy analysis (dbrDA) to estimate the effects of each explanatory variables<sup>44</sup>. We performed dbrDA using Bray–Curtis dissimilarity distance separately for each study site using the same variables at each site. Our dbrDA produced a total of four constrained axes and 29 unconstrained axes, the first two constrained axis (referred to as canonical axis, CAP) are shown in Fig. 4. The results show that on Phra Thong Island, the marine sediment samples, intertidal and beach sands clustered separately (16% of the variance) from the backdune pit samples and swale core samples (canonical axis 1 in Fig. 4a; ANOVA  $F_{1,15} = 4.0117$ ,  $p$ -value = 0.001 in Supplementary Table 3). A similar cluster was observed in Cuddalore samples, where the marine sediment samples and beach sands were clustered separately from the sandy backdune pit samples, supported with 25.8% of the variance (canonical axis 1 in Fig. 4b; ANOVA  $F_{1,29} = 13.4769$ ,  $p$ -value = 0.001 in Supplementary Table 3).

As we focus the analysis to overwash deposits and the intercalating sediment samples, we observed that on Phra Thong Island, the first canonical axis (19.2% of the variance) separated the 2007 storm deposits from the 2004 IOT deposits and terrestrial samples (Fig. 4c; ANOVA  $F_{1,11} = 2.8694$ ,  $p$ -value = 0.003 in Supplementary Table 3) while the 2004 IOT deposits were clustered separately from other sample types along the second canonical axis (8.6% of the variance; Fig. 4c; ANOVA  $F_{1,11} = 1.2954$ ,  $p$ -value = 0.056 in Supplementary Table 3). Multivariate analyses of community structure and composition for Phra Thong Island samples revealed that all deposit types differ significantly (PERMANOVA composition: pseudo- $F_{4,13} = 2.4865$ ,  $p < 0.0001$  in Supplementary Table 4). Post-hoc pairwise comparisons show that the 2007 storm deposit microbial communities differed from all other sample types and beach sediments differed from the soil and terrestrial samples (see Data Availability). The 2004 IOT deposits did not differ from the soil and terrestrial samples and marine sediments despite being <1% similar to marine

sediments and <9% similar to soil and terrestrial samples (Supplementary Fig. 6).

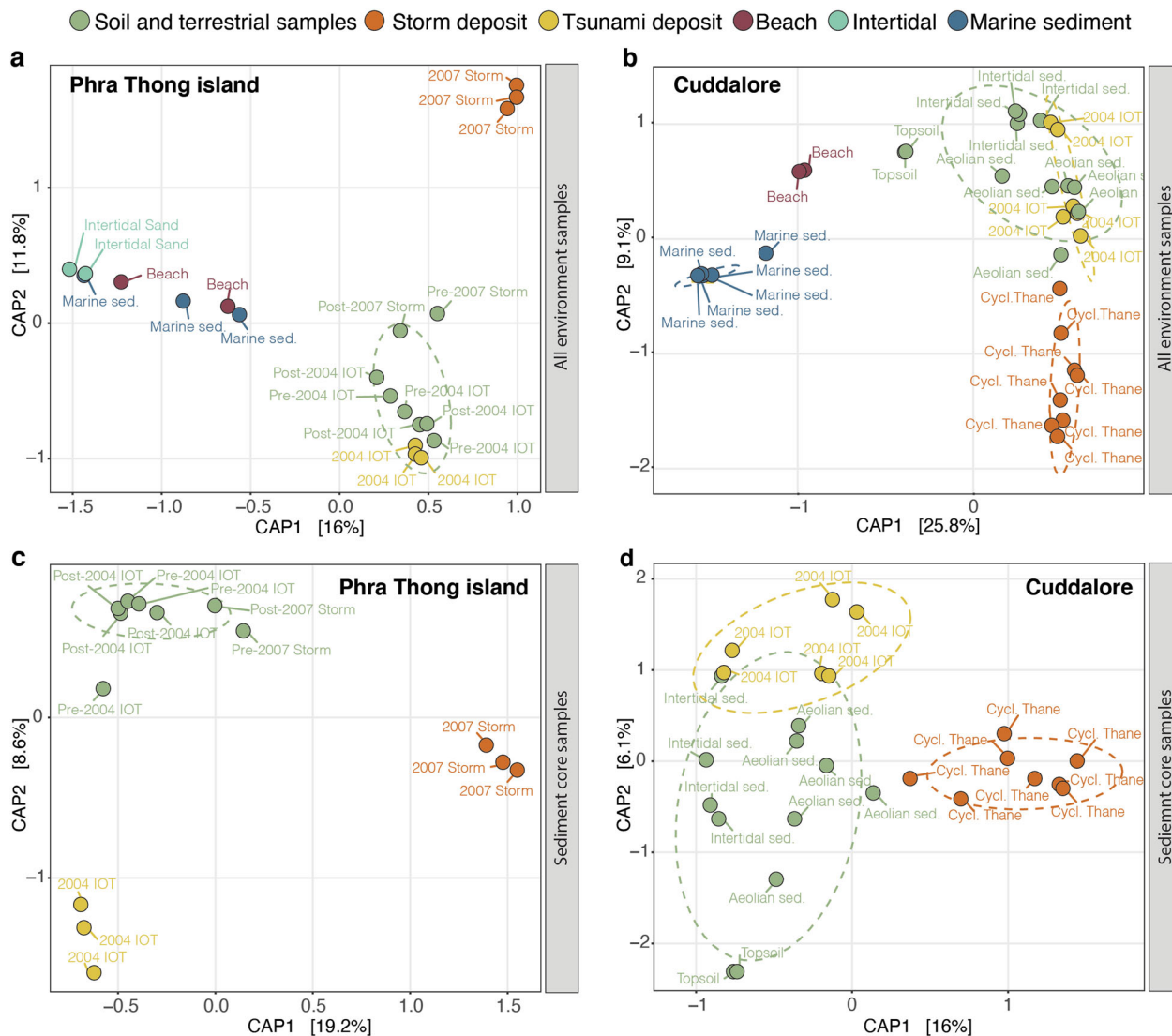
A similar result was observed in Cuddalore. Cyclone Thane deposits were significantly different from the 2004 IOT, and overlying and underlying sediments, supported with 16% of the variance (canonical axis 1 in Fig. 4d; ANOVA  $F_{1,23} = 4.5135$ ,  $p$ -value = 0.001; Supplementary Table 3). The second canonical axis in Fig. 4d marginally separates the 2004 IOT deposits from the intercalating samples supported with 6.1% of the variance (ANOVA  $F_{1,23} = 1.6988$ ,  $p$ -value = 0.081, Supplementary Table 3). Nonetheless, multivariate analyses on community composition for samples from Cuddalore (Supplementary Table 4) showed that all sample types differed significantly from each other, except for tsunami deposits and the overlying and underlying aeolian and reworked sediments (PERMANOVA composition: pseudo- $F_{4,29} = 5.674$ ,  $p < 0.0001$ ). It must be noted that the lower dissimilarity between tsunami deposits and intercalating soil may be due to the significantly higher dispersion in aeolian deposits (>50%) than marine sediment, beach, and overwash deposits (see Supplementary Material). On average we find that the tsunami deposits were <30% similar to storm deposits and <25% similar to aeolian deposits (Supplementary Fig. 6).

**Temporal variability of microbial dissimilarity between sample type.** We sampled the deposits preserved on Phra Thong Island in consecutive years (2014 and 2015) at adjacent sites (<1 m away). That data showed that the differences in microbial community composition among sample types were consistent through time (type by year interaction:  $F_{2,13} = 0.8520$ ,  $p = 0.7337$ ; year:  $F_{1,13} = 1.0434$ ,  $p = 0.4236$ ; Supplementary Table 4). The microbial population remained relatively unchanged despite our sampling sites being in a tropical setting with frequent rainfall and significant fluctuation of groundwater levels<sup>5,9,16,17</sup>.

**Characterization of the deposition source with specific microbial markers.** Differential analysis using a negative binomial generalized linear models (*DESeq2* package; 43) revealed that within the 527 ASVs identified on Phra Thong Island samples and 92 ASVs identified in Cuddalore samples (Fig. 5 - left dendrogram), there were substantial differences in the microbial communities preserved in the storm deposits and the 2004 IOT deposits. At both study sites, storm deposits were clustered together (except one storm deposits labeled as 77cmS2 at Cuddalore), and separated from tsunami and intercalating sediment samples (Fig. 5—top dendrogram). Likewise, tsunami deposits on Phra Thong Island are clustered together (Fig. 5a—top dendrogram). Nonetheless, the differences between tsunami and intercalating aeolian sediments at Cuddalore are less clearly defined (Fig. 5b—top dendrogram).

In general, ASVs that were prominent in the 2007 storm and Cyclone Thane deposits were from the class Gammaproteobacteria and class Actinobacteria (Fig. 5). ASVs that were unique to the 2007 storm deposits were from the family Blastocatellaceae (Subgroup 4), family Subgroup 6, family Holophagae, family Nitrososphaeraceae, family Nitrosoomonadaceae, family Microscillaceae, and family Acetobacteraceae (Fig. 5a). ASVs that were unique to the Cyclone Thane deposits were from the family Parcubacteria, family Chromobacteriaceae, family Rubinisphaeraceae, family Burkholderiaceae, family Micromonosporaceae, family Bacillaceae, family Nocardioideaceae, family Sporichthyaceae, family Caulobacteraceae and family Sericytochromatia (Fig. 5b). Notably, family Chitinophagaceae and family Thermoplasmata appeared to be taxa that were unique across both storm deposits (Fig. 5) and not found in other samples at each site. There were no taxa that were only present within the Phra Thong





**Fig. 4 Distance-based redundancy analysis (dbRDA) using Bray-Curtis dissimilarity.** dbRDA shows that microbial communities in marine sediments, beach, and intertidal samples were different from backdune pit and swale core samples (a, b). c and d show that at Phra Thong Island and Cuddalore, storm deposits and 2004 IOT deposits have distinct microbial communities that are dissimilar to soil and terrestrial samples. The ellipses (dashed line) bound each cluster by sample type at an 80% confidence interval. The significant values of respective canonical axes are reported in Supplementary Table 2 and calculated based on multivariate analysis with 9999 permutations.

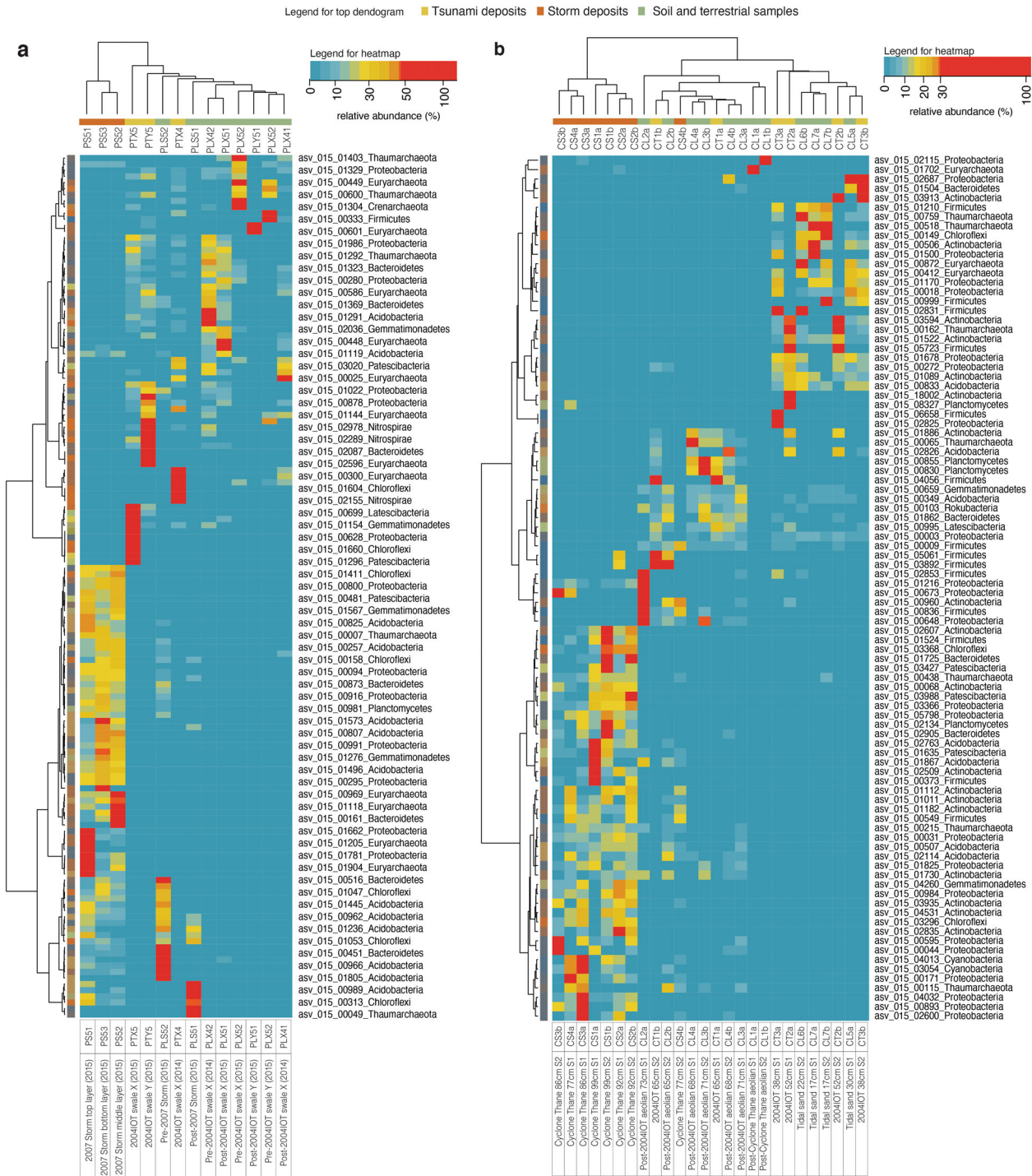
Island 2004 IOT deposits. Nonetheless, we did identify family Clostridia and family Paenibacillaceae ASVs that were unique to the Cuddalore 2004 IOT deposits.

**Discussion**

**Do microbial communities differ between the respective study sites and samples?** Our study shows that the microbial communities from the Phra Thong Island and Cuddalore samples were significantly different from each other and that these differences also existed between the microbial communities in the offshore, onshore, backdune and swale environments (Supplementary Fig. 3). We also observed that the microbial communities preserved on Phra Thong Island were significantly dissimilar to those at Cuddalore (Supplementary Fig. 3). These large differences in community composition likely correspond to differences in the site geomorphology and different sedimentary processes that build the sedimentary profiles. Phra Thong Island has organic-rich swales with fine to medium-grained sandy beach ridges<sup>5,16</sup>, while Cuddalore’s sediment is mainly sand with limited

differences in grain size distribution and characteristically low organic and carbonate content<sup>39</sup> (Fig. 1, Supplementary Fig. 1). The microbial community structure, diversity, and distribution at both sites are primarily affected by sediment physicochemical characteristics followed by geomorphological characteristics such as grain size and sediment type<sup>45</sup>.

**Is it possible to identify individual tsunami and storm deposits from overlying and underlying soils and sediments?** Our first focus was to investigate the possibility that microbial meta-barcoding approaches can identify individual tsunami and storm deposits from the overlying and underlying soils or sediments. First, we examined the storm deposits microbial communities, and the result shows that our approach can distinguish the 2007 storm deposit on Phra Thong Island and Cyclone Thane deposits at Cuddalore from the soil and sediments ( $p$ -value = 0.0001; Fig. 4, Supplementary Table 4). Notably, our approach is capable of identifying sediment deposited from relatively minor storm events. This finding highlights the clear potential for



**Fig. 5 Heatmap highlighting the unique ASVs that differentiate the 2004 IOT and a storm deposit that impacted the same location.** Heatmap plotted using output from differential analysis selected using a threshold  $p$ -value  $< 0.001$ . This plot shows the relative abundance (%) of members of the microbial community that contributed significantly to differentiating between 2004 IOT and the 2007 storm deposits on Phra Thong Island (**a**) and between 2004 IOT and Cyclone Thane deposits at Cuddalore (**b**). The top dendrogram indicate the clustering of tsunami deposits (yellow), storm deposits (orange) and soil and terrestrial samples (green). Detailed sample description of the top dendrogram is presented at the bottom of the panel. ASVs were labeled on the right side of the heatmap with taxonomic grouping at phylum level.

applying microbial approaches to aid historical coastal hazard study on identifying tsunami and storm deposits.

The next research question examined whether it is possible to identify/differentiate 2004 IOT tsunami deposits from the soil and sediments above and below it. The microbial community dissimilarity between tsunami deposits and soil and terrestrial

samples is significant ( $p$ -value  $< 0.0001$ ; Fig. 4, Supplementary Table 4). Our observation is in agreement with the studies summarized in Table 1 that microbial communities were different between tsunami deposited sediments and non-tsunami soils and sediments. Similar to Nayak et al.<sup>25</sup>, Asano et al.<sup>28</sup>, Godson et al.<sup>26</sup>, and Somboonna et al.<sup>29</sup>, we observed that tsunami

deposits have relatively lower diversity as compared to non-tsunami soils and sediments (Fig. 3). Likewise, Ramesh et al.<sup>24</sup> reported that marine sediment has higher microbial diversity immediately after a tsunami flooding event and the microbial community diversity gradually decreased in subsequent days. We hypothesize that an overwash deposit transported from offshore to onshore could transport material from various environments, and thus, contains a richer microbial diversity. The abrupt changes during the flooding event, and the recovery of the landscape to the pre-event condition may have reduced the diversity over time. Nonetheless, our 2014 and 2015 sampling in Phra Thong Island suggests that during a normal year, there is little microbial variation in any of the samples (Supplementary Table 4). Hiraoka et al.<sup>27</sup> is the only study reported that tsunami flooded soil contains a rich microbial community as compared to unaffected soil, however, they are only referring to one genus that was successfully isolated from the samples.

When compared to the storm deposits, the microbial signature in the Cuddalore 2004 IOT deposits was more subtle as the tsunami deposits have about 30% similarity with the overlying and underlying soil layers (Supplementary Fig. 6). We observed that microbial communities in aeolian sediments post-2004 IOT are grouped with microbial communities in the upper part of the 2004 IOT deposits. Meanwhile, the microbial communities from the lower portion of the 2004 IOT deposit are grouped with tidal sediment pre-2004 IOT (Fig. 5b). Our observation may suggest mixing of underlying sediments into the 2004 IOT deposits during deposition followed by a subsequent mixing of tsunami deposits in the overlying reworked sands. The mixture between sediment units occurred as a tsunami inundation may erode material from the underlying substrate or land surface<sup>46</sup>. It could also affect the incorporation of underlying sediments during tsunami backflow to the sea<sup>4</sup>.

#### **Can we distinguish between modern storm and 2004 IOT deposits based on ASVs?**

When we compare modern storm and 2004 IOT deposits based on ASVs, our microbial metabarcoding approach can reliably distinguish between tsunami and storm deposits. Our multivariate analysis shows that the differences in microbial communities between tsunami and storm deposits, as well as in soil and terrestrial and offshore sediments, were statistically significant (Fig. 4; Supplementary Table 4). Therefore, our approach can diagnose an overwash event that is different from typical soil microbial composition for both sites. Tsunami and storm have distinct sedimentary deposition mechanisms<sup>45</sup> but both extreme waves can transport sediment from the near-shore marine environment<sup>47,48</sup>. Some tsunami waves have been reported to contain offshore sediments from inner continental shelf environments<sup>49,50</sup> that were scoured from 30 to 50 m water depth. It is likely that such sediment would contain distinctive microbial taxa that are unlikely to be found in the intertidal and beach sediments that dominate the source of storm deposits, along the same coastline. Furthermore, all locations may be modified by post-event weather conditions including strong winds and extreme precipitation that can potentially alter the soil microbial communities<sup>51</sup> either directly (e.g. transporting unique microbial communities) or indirectly (e.g. altering the soil chemistry)<sup>52,53</sup>.

#### **Can we use a eukaryote community-level approach to identify overwash deposits?**

We extended our metabarcoding approach to include the 18S rRNA gene, which is a universal genetic marker for eukaryotes<sup>54</sup>. The analysis of eukaryotic community shows no significant differences in discriminating storm, tsunami, and soil and terrestrial samples (Supplementary Fig. 4). This may be due

to the low efficiency in our primer sets to amplify eukaryotic rRNA gene, as eukaryotic rRNA genes tend to have long variable region, gene region that demonstrates considerable sequence diversity among different eukaryotes<sup>54</sup>, while next-generation sequencing technology has restricted sequencing length. Instead of using a community-level approach, we suggest a metabarcoding approach targeting a specific group of eukaryotes.

**Can we identify the source of the overwash deposits?** Since we can identify overwash deposits from intercalating soils and sediments, the next challenge is to identify the source of the overwash deposit. Our study found no similarity between 2004 IOT deposits and marine sediment (Supplementary Fig. 6). Signature taxa that can identify storm and tsunami deposits live in a vast range of habitats. This result possibly suggests that microorganisms with the capacity to tolerate a broad range of environmental conditions increased in relative abundance after the pulse disturbance of the overwash as opposed to “habitat specialists”<sup>55</sup>, that would have found themselves at a competitive disadvantage. Our finding is in line with other tsunami microbial studies (Table 1) who report microbial community changed before and after a tsunami inundation event, and these community changes after the flooding disturbance remained up to 10 years after the event<sup>21</sup>.

#### **Is there a global microbial signature for overwash deposits?**

We identified 527 ASVs on Phra Thong Island and 92 ASVs at Cuddalore that signify the differences between tsunami and storm deposits despite the ubiquity of most of the taxa found in this study. In particular, the family Chitinophagaceae and the family Thermoplasmata are found only in storm deposits at both locations. While we find this a promising step in the search for definitive signatures, we have limited evidence to conclude that these taxa are a global signature for storm deposits. Furthermore, we did not identify any ASVs that are only present in the 2004 IOT deposit at either location.

Our work addressed six primary questions on the use of microbes to investigate extreme wave events. Earlier attempts to use microbial communities to identify tsunami deposits (Table 1) commonly suffered from limitations in identifying and characterizing the microbial communities due to methodological restrictions. The experimental design in these earlier studies was also not able to adequately examine the potential of using microbial communities to tackle the struggle in historical coastal hazard study—to identify overwash deposit from the geological records. Our work overcomes these limitations by applying advanced molecular techniques and robust statistical testing. The microbial metabarcoding approach that we adopt clearly demonstrates that microbial communities do differ between recent tsunami and storm deposits preserved in similar geomorphic settings at two distinct locations. We acknowledge that our study focuses only on modern tsunami and storm flooding events, and thus it will be interesting to investigate older or palaeo-overwash deposits from other sites. Such work will facilitate an examination of the potential of microbial signatures existing in overwash deposits over long time frames. Nevertheless, we concede that our work could not confidently identify the factor/s that caused the community structure differences between tsunami and storm deposits. Future work involving detailed chemical analysis such as oxygen and heavy metal ions present in the sediments will facilitate the understanding of what caused the changes and how microbes respond to disturbance induced by coastal flooding. We present for the first time, unambiguous evidence for discriminating tsunami and storm sedimentary deposits using the metabarcoding approach. Our study addresses

a key challenge in the analysis of coastal hazards that could support improved risk assessment for coastal regions.

## Methods

**Study site description and sample collection.** Phra Thong Island is approximately 125 km north of Phuket on the west coast of southern Thailand in the Andaman Sea (Fig. 1a, b(i), b(ii)). This island was impacted by the 2004 Indian Ocean Tsunami (2004 IOT) when the Sumatran megathrust ruptured. Since the formation of beach ridge sequences and a series of swales on the western side of Phra Thong Island are geomorphologically favorable for preserving tsunami deposits<sup>5,16</sup>, the site was used in over 10 studies examining tsunami events<sup>9</sup>. The 2004 IOT deposits are preserved in the marshy swale—Swale X (9° 8.1742' N, 98° 15.916' E) and Swale Y (9° 7.917' N, 98° 15.744' E) in Fig. 1b(ii), c, d; Supplementary Data 1), underneath an organic soil layer<sup>5</sup>. We obtained local permission from Chulalongkorn University to conduct research at this site and received approval from the land-owner to perform sample collection. To avoid cross-contamination between samples and interference of modern DNA to the targeted samples, all tools such as coring tube, Van Veen grab, hacksaw, hand trowel and Van Dorn bottle were pretreated with 20% bleach solution and rinsed between samples. Powder-free surgical gloves were worn when handling the samples. Sediment cores from Swale X and Swale Y were collected using a plastic push-core that was pretreated with a 20% bleach solution. Each stratigraphic layer was subsampled using a sterile 50 mL conical tube. Six sediment samples were collected from Swale X, three sediment samples were collected from Swale Y and 5 sediment samples were collected from a backdune pit (Supplementary Data 1). Aside from collecting sediment samples from the geological record, we also collected environmental samples; beach and intertidal sediments using sterile 50 mL conical tube, as well as marine sediments from 2m and 20m water depths using a Van Veen grab (Fig. 1 b(i), b(ii), d; Supplementary Data 1). We used a Van Dorn bottle to collect marine water from 2m and 20m water depth. The sedimentary profile, i.e. grain size characteristics and mineralogy of the environmental samples, varies between offshore, onshore sediment and overwash deposits<sup>16</sup>.

Our study site in India is located along the coastline of Cuddalore, Tamil Nadu, southeast India (Fig. 1a, e(i), e(ii)). The southern part of India was affected by the 2004 IOT, with the coast of Tamil Nadu experiencing the highest death toll and most severe damage<sup>56</sup>. Anna University provided permission and field assistance in sample collection from the study site. We pretreated all the tools with 20% bleach solution, rinsed the tools in between collecting each sample to prevent cross-contamination between samples. We wore dust-free surgical gloves during sampling to avoid the potential introduction of modern DNA into the samples. We collected 30 sediment samples from a backdune pit in Devanampattinam, Cuddalore (11° 45.198' N 79° 47.334' E; Fig. 1e(i), e(ii); Supplementary Data 1), as they contain both the 2004 IOT deposits and the 2011 Cyclone Thane deposits<sup>38–40</sup>. The backdune pit is made up primarily of quartz with little organic and carbonate content. Underneath the surface aeolian sediment is the 2011 Cyclone Thane deposit, and below the storm deposits is reworked aeolian sand followed by the 2004 IOT deposits (Fig. 1f, g; Supplementary Data 1). Beneath the 2004 IOT deposits is a fine- to medium-grained sand unit laminated with heavy minerals (Fig. 1f, g; Supplementary Data 1). We also collected sediment samples from the beach, a nearby lagoon that is rich in organic matter, and marine sediment samples (at 5 m, 9.5 m, and 15 m water depths; Fig. 1e(i), g; Supplementary Data 1) using a Van Veen grab pretreated with 20% bleach solution. All samples were collected using sterile 15 mL conical tubes. A Van Dorn bottle pretreated with 20% bleach solution is used to collect lagoon and marine water samples. The water samples were filtered through a 0.20 µm sterivex filter to capture the water community. Upon collection, all the samples (both from Phra Thong Island and Cuddalore) were transported using a portable liquid nitrogen dry shipper and stored in the Ultra-Low Freezer facility at -80 °C at the Asian School of the Environment, Nanyang Technological University, Singapore immediately upon arrival.

**Grain-size analysis.** Each overwash deposits and its overlying and underlying samples were treated with 15% Hydrogen Peroxide (H<sub>2</sub>O<sub>2</sub>). This chemical treatment is to remove organic materials that are present in the samples to determine the granulometry of the clastic component. The samples were washed with deionized water at least three times before performing grain size analysis using the Malvern Mastersizer 3000. The Malvern Mastersizer 3000 uses laser obscuration to measure particle size distribution. Each sample was subjected to 1 min sonication to disaggregate the sediment before taking the measurement. The raw data was analyzed using GRADISTAT version 9.1<sup>57</sup>. This program calculates grain size statistic using the Folk and Ward and moments method<sup>58</sup>, and generate statistics including the mean(phi) and sorting that were reported in Fig. 1.

**Chemical analysis.** Dried bulk sediment samples were sent to the Stable Isotope Laboratory at Hong Kong University to be analyzed for TOC, TN, and TS. Briefly, 30 mg of sediment samples were weighed in a 5 mm × 9 mm silver capsule (Sercon) and directly acidified with 6 N of hydrogen chloride (HCl) to remove residual carbonate before analysis. All samples were dried overnight at 60 °C and

combusted and analyzed using an Elemental Analyzer attached to an Isotope Ratio Mass Spectrometer (EA-IRMS).

**DNA extraction from sediments.** We extracted total DNA from 250 mg sediment samples using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany). We started the extraction by adding 200 µL of Phenol: Chloroform: Isoamyl-alcohol (25:24:1)<sup>59</sup> before adding Solution C1. After centrifuging at 28,000g for 1 min, we transferred the supernatant to a new tube and added 100 µL Solution C2 and 100 µL Solution C3. The mixture is briefly vortexed and incubated at 4 °C for 5 min. We then centrifuged the mixture at 28,000g for 1 min and transferred the supernatant to a new tube, adding Solution C4 in the ratio of 1:1 with the supernatant, and 650 µL of 100% molecular grade ethanol. For the remaining steps, we followed the manufacturer's protocol. The extracted DNA was cleaned with OneStep™ PCR Inhibitor Removal Kit (ZYMO Bioresearch, Irvine, CA) to remove any potential impurities such as polyphenols, humic acids, and flavic acids from the soil that can inhibit downstream amplification reactions. The DNA concentration was determined using the Qubit® Fluorometer (Thermo Fisher Scientific, USA).

**Construction of amplicon sequencing library.** We performed a polymerase chain reaction (PCR) on the extracted total DNA to amplify a targeted fragment from the genetic material using a short single-stranded DNA primer set. We used two different primers sets: 926wF (AAA CTY AAA KGA ATT GRC GG) and 1392R (ACG GGC GGT GTG TRC) is a universal primer that amplifies the 16S rRNA gene V6-V8 regions in archaea, bacteria and eukaryotes<sup>60,61</sup>; while 574F (GCG GTA ATT CCA GCT CCA A) and 1132R (CCG TCAA TTH CTT YAA RT) amplifies the 18S rRNA gene V4-V7 regions in eukaryotes<sup>54</sup>. Each reaction is comprised of 12.5 µL of 2x KAPA HiFi Hotstart Ready Mix (KAPA Biosystems, Cape Town, South Africa), 10 µL of 1 µM of primer, and 2.5 µL of 5 ng/ml of total DNA, with each sample done in triplicate. A PCR consists of three steps: 1. separation of the double-stranded DNA into single strands at high temperatures (denaturation), 2. binding of the DNA primers to the edge of the targeted region (annealing) 3. building the complementary strand of the targeted fragment through polymerization (extension). We started by denaturing the double-stranded DNA at 95 °C for 3 min, then steps (1) through (3) for 20 times: (1) denaturation at 95 °C for 30 s, (2) primer annealing at 55 °C for 30 s, and (3) extension of the complementary strand at 72 °C for 30 s. This is followed by a final extension at 72 °C for 5 min. The PCR was performed in a ThermoFisher SimpliAmp Thermal Cycler. To prevent any potential bias in this PCR-based method, we used the minimum amplification cycle and maintained the same PCR setting and reaction mixture for all the samples. All the PCR products were purified with Agencourt AMPure XP beads (Beckman Coulter, Singapore) following the manufacturer's protocol. These PCR products were sent to the Macrogen Asia Pacific Pte. Ltd. sequencing facility to generate millions of DNA sequences using an Illumina MiSeq machine with MiSeq Reagent Kit v3 chemistry (2 × 300 bp).

**Sequence analysis.** The raw sequences generated from Illumina MiSeq were initially processed by removing primers from the sequences using cutadapt version 2.10<sup>62</sup>. Next, the sequences were filtered by truncating those that are shorter than 280 in forward reads (R1), and 230 in reverse reads (R2). We also removed sequences that have a quality score of less than or equal to 2 and have an expected error rate (sum (10<sup>-quality score</sup>/10)) of higher than 2 (R1) and 5 (R2) using the filterAndTrim function in DADA2 package in R. We applied DADA2<sup>63</sup> that used a novel algorithm to calculate the error introduced during sequencing to construct an ASVs table. This ASV is analogous to the traditional operational taxonomic units (OTUs), it infers the true sample composition as oppose to OTUs that cluster sequences with a fixed dissimilarity threshold i.e. 97% similarity<sup>64</sup>. Taxonomy classification was performed using SILVA version 132 database<sup>65</sup> as a reference to identify bacteria and archaea for the 16S rRNA dataset, and PR2 database version 4.12.0<sup>66</sup> (<https://github.com/pr2data-base/pr2database>) as a reference to identify eukaryote using RDP naïve Bayesian classifier as implemented in R DADA2 package for the 18S rRNA dataset. ASVs that have less than 10 counts were removed, followed with a bootstrap value lower than 90% at the supergroup/ phylum level were discarded. The final ASV table contained a total number of 25,034 ASVs generated from 926F-1392R primers and 5840 ASVs from 574F-1132R primers.

**Statistical analysis.** Statistical analysis was performed in R version 3.6.1<sup>67</sup> using the *phyloseq* package. All relevant R code is available from the associated GitHub Repository (see section Data and Code availability). The number of reads in each sample was normalized using median sequencing depth. The richness of the samples was calculated using the Shannon index (H') and Simpson index (D) available in the *vegan* package<sup>68</sup> to accounts for both ASVs abundance and evenness in the dataset. These samples were rarefied to the minimum number of ASVs across all samples to simulate an even number of reads per sample before calculating the ASVs richness. Ordination analysis was performed to understand how the community differs from one sample from the other. The number of ASVs was transformed with square-root to minimize the effect of abundances

toward composition analyses. We applied PCoA calculated using the Bray–Curtis dissimilarity index available in the *vegan* package to measure the similarity and dissimilarity between samples type. We included the result from the Envfit analysis (also available in the *vegan* package)<sup>68</sup> into the PCoA plot to understand whether the dissimilarity between different samples type is driven by changes in the environment chemistry. The Envfit result was calculated with 9999 permutations generated using TOC, TN, and TS. In subsequent analysis, we separated the dataset based on the study site and removed water samples and lagoon sediment samples from the dataset to standardize the testing variables. No water samples and lagoon sediment samples were collected from Phra Thong Island. We performed a constrained ordination with distance-based redundancy analysis (dbRDA) that is available in the *phyloseq* package ordinate function to visualize the dissimilarity between sample types and their driving explanatory variables. The ellipses included in the dbRDA plots were calculated with 80% confidence intervals.

We used dissimilarity-based permutational multivariate analyses of variance (PERMANOVA<sup>69</sup>) to examine differences in microbial community structure (square-root transformed, normalized median sequences) between sample types in each location. Analyses were undertaken in PRIMER v7 (PRIMER-E, UK). Bray–Curtis similarities were calculated for all pairs of samples. Sample type was a fixed factor, with five levels. For the Phra Thong Island dataset, we also included year as a fixed, orthogonal factor (two levels: 2014, 2015) to determine whether the effects of sample type were consistent in time. *p*-values were calculated using 9999 unrestricted permutations of raw data (Cuddalore; appropriate for one-factor analyses) or restricted permutations of residuals (Phra Thong Island; appropriate for multi-factorial analyses<sup>70</sup>). A posteriori pairwise comparisons were done to determine which sample types differ from which, with *p*-values calculated using Monte Carlo simulations. Permutational multivariate analyses of dispersion were used to examine differences in dispersion among sample types at each location. *p*-values were calculated using 9999 permutations.

We used the *DESeq2* package<sup>71</sup> to perform differential analysis to determine which ASVs respond to the differences between 2004 IOT deposits and storm deposits at the respective study sites. *DESeq2* is a negative binomial generalized linear models and uses the Wald test for significance testing. An adjusted *p*-value (*p*-value corrected for multiple hypothesis testing<sup>72</sup>) of 0.001 was used as the cutoff to select the ASVs that were significantly different present in either 2004 IOT deposits or storm. The differentially present ASVs were extracted from the dataset and visualized using a heatmap. A hierarchical cluster analysis was performed on these ASVs using Euclidean distance to group ASVs based on the similarity between sample type using Ward's minimum variance method<sup>73</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Raw sequence data that support the finding of this study have been deposited in National Center for Biotechnology Information (NCBI) GenBank under the BioProject ID [PRJNA5433068](https://doi.org/10.21969/ncbi/bioproject/5433068). All statistical results are reported in <https://github.com/slimelab/Tsunami-microbes>.

## Code availability

All relevant R code is available from the GitHub repository: <https://github.com/slimelab/Tsunami-microbes>.

Received: 27 October 2020; Accepted: 26 May 2021;

Published online: 18 June 2021

## References

- Nicholls, R. J. et al. in *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) Ch. 6 (Cambridge University Press, 2007).
- Gordon, M. et al. in *Global Assessment Report on Disaster Risk Reduction* Ch. 3 (UNDRR, 2019).
- Dominey-Howes, D. Documentary and geological records of tsunamis in the Aegean Sea region of Greece and their potential value to risk assessment and disaster management. *Nat. Hazards* **25**, 195–224 (2002).
- Switzer, A. D., Yu, F., Gouramanis, C., Soria, J. & Pham, T. D. Integrated different records to assess coastal hazards at multi-century timescales. *J. Coastal Res.* **70**, 723–728 (2014).
- Jankaew, K. et al. Medieval forewarning of the 2004 Indian Ocean tsunami in Thailand. *Nature* **455**, 1228–1231 (2008).
- Liu, K. B. & Fearn, M. L. Reconstruction of prehistoric landfall frequencies of catastrophic hurricanes in northwestern Florida from lake sediment records. *Quaternary Res.* **54**, 238–245 (2000).
- Donnelly, J. P. & Woodruff, J. D. Intense hurricane activity over the past 5,000 years controlled by El Niño and the West African monsoon. *Nature* **447**, 465–468 (2007).
- Nanayama, F. et al. Unusually large earthquakes inferred from tsunami deposits along the Kuril trench. *Nature* **424**, 660–663 (2003).
- Gouramanis, C. et al. High-frequency coastal overwash deposits from Phra Thong Island, Thailand. *Sci. Rep.* **7**, 1–9 (2017).
- Nanayama, F. et al. differences between the 1993 Hokkaido-nansei-oki tsunami and the 1959 Miyakojima typhoon at Taisei, southwestern Hokkaido, northern Japan. *Sediment. Geol.* **135**, 255–264 (2000).
- Morton, R. A., Gelfenbaum, G. & Jaffe, B. E. Physical criteria for distinguishing sandy tsunami and storm deposits using modern examples. *Sediment. Geol.* **200**, 184–207 (2007).
- Marriner, N. et al. Tsunamis in the geological record: Making waves with a cautionary tale from the Mediterranean. *Sci. Adv.* **3**, e1700485 (2017).
- Vött, A. et al. Returning to facts: response to the refusal of tsunami traces in the ancient harbour of Lechaion (Gulf of Corinth, Greece) by ‘non-catastrophists’ – Reaffirmed evidence of harbour destruction by historical earthquakes and tsunamis in AD 69–79 and the 6<sup>th</sup> cent. AD and a preceding pre-historical event in the early 8<sup>th</sup> cent. BC. *Zeitschrift Geomorphologie* **61**, 275–302 (2018).
- Shanmugam, G. The tsunamite problem. *J. Sediment. Res.* **76**, 718–730 (2006).
- Chagué-Goff, C., Chan, J. C. H., Goff, J. & Gadd, P. Late Holocene record of environmental changes, cyclones and tsunamis in a coastal lake, Mangaia, Cook Islands. *Isl. Arc* **25**, 333–349 (2016).
- Pham, D. T. et al. Elemental and mineralogical analysis of marine and coastal sediments from Phra Thong Island, Thailand: Insights into the provenance of coastal hazard deposits. *Mar. Geol.* **385**, 274–292 (2017).
- Sawai, Y. et al. Diatom assemblages in tsunami deposits associated with the 2004 Indian Ocean Tsunami at Phra Thong Island, Thailand. *Mar. Micropaleontol.* **73**, 70–79 (2009).
- Pilarczyk, J. E. et al. Microfossils from coastal environments as indicators of paleo-earthquakes, tsunamis and storms. *Palaeogeogr. Palaeoclimatol.* **413**, 144–157 (2017).
- Gouramanis C. in *Geological Records of Tsunamis and other Extreme Waves* (eds Engel, M., Pilarczyk, J., May, S. M., Brill, D. & Garrett, E.) Ch. 13 (Elsevier, 2020).
- Goff, J., Chagué-Goff, C., Nichol, S., Jaffe, B. & Dominey-Howes, D. Progress in palaeotsunami research. *Sediment. Geol.* **243**, 70–88 (2012).
- Asano, R. et al. Changes in bacterial communities in seawater-flooded soil in the four years after the 2011 Tohoku tsunami in Japan. *J. Mar. Sci. Eng.* **8**, 76 (2020).
- Atwater, B. F. et al. Extreme waves in the British Virgin Islands during the last centuries before 1500 CE. *Geosphere* **13**, 301–368 (2017).
- Jentsch, A. & White, P. A theory of pulse dynamics and disturbance in ecology. *Ecology* **100**, e02734 (2019).
- Ramesh, S., Jayaprakashvel, M. & Mathivanan, N. Microbial status in seawater and coastal sediment during pre- and post-tsunami periods in the Bay of Bengal, India. *Mar. Ecol.* **27**, 198–203 (2006).
- Nayak, A. K. et al. Post tsunami changes in soil properties of Andaman Islands, India. *Environ. Monit. Assess.* **170**, 185–193 (2010).
- Godson, P. S., Chandrasekar, N., Kumar, S. K. & Vimi, P. V. Microbial diversity in coastal sediments during pre- and post-tsunami periods in the south east coast of India. *Front. Biol.* **9**, 161–167 (2014).
- Hiraoka, S. et al. Genomic and metagenomics analysis of microbes in a soil environment affected by the 2011 Great East Japan Earthquake tsunami. *BMC Genomics* **17**, 1–13 (2016).
- Asano, R. et al. Seawater inundation from the 2011 Tohoku Tsunami continues to strongly affect soil bacterial communities 1 year later. *Microb. Ecol.* **66**, 639–646 (2013).
- Somboonna, N. et al. Microbial ecology of Thailand tsunami and non-tsunami affected terrestrials. *PLoS ONE* **9**, e94236 (2014).
- Tas, N. et al. Impact of fire on active layer and permafrost microbial communities and metagenomes in an upland Alaskan boreal forest. *ISME J* **8**, 1904–1919 (2014).
- Dooley, S. R. & Treseder, K. K. The effect of fire on microbial biomass: a meta-analysis of field studies. *Biogeochemistry* **109**, 49–61 (2012).
- Kawagucci, S. et al. Disturbance of deep-sea environments induced by the M9.0 Tohoku Earthquake. *Sci. Rep.* **2**, 1–7 (2012).
- Morimura, S., Zeng, X., Noboru, N. & Hosono, T. Changes to the microbial communities within groundwater in response to a large crustal earthquake in Kumamoto, southern Japan. *J. Hydrol.* **581**, 124341 (2020).
- Olsen, G. J., Lane, D. J., Giovannoni, S. J. & Pace, N. R. Microbial ecology and evolution: a ribosomal RNA approach. *Annu. Rev. Microbiol.* **40**, 337–365 (1986).

35. Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**, 669–685 (2004).
36. Szczuciński, W. et al. Ancient sedimentary DNA reveals past tsunami deposits. *Mar. Geol.* **381**, 29–33 (2016).
37. Neelson, K. H. Sediment bacteria: who's there, what are they doing, and what's new? *Annu. Rev. Earth Pl. Sc.* **25**, 403–434 (1997).
38. Srinivasalu, S., Karthikeyan, A., Switzer, A. D. & Gouramanis, C. Sedimentological characteristics of tsunami and storm deposits: a modern analogue from Southeast Indian Coast. In *Paper Presented at the AOGS-AGU Joint Assembly, Singapore, 13–17 September 2012* (2012).
39. Switzer, A. D., Srinivasalu, S., Thangadurai, N. & Mohan, V. R. Bedding structures in Indian tsunami deposits provide clues to the dynamics of tsunami inundation. *Geol. Soc. Spec. Publ.* **361**, 61–77 (2012).
40. Gouramanis, C. et al. Same Same, but different: sedimentological comparison of recent storm and Tsunami deposits from the south-eastern coastline of India. In *Paper presented in AGU Fall Meeting (NH21A-3811), San Francisco, California, 15 – 19 December 2014* (2014).
41. Fisher, R. A., Corbet, A. S. & Williams, C. B. The relation between the number of species and the number of individuals in a random sample of animal population. *J. Anim. Ecol.* **12**, 42–58 (1943).
42. Hurlbert, S. H. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577–586 (1971).
43. Xu, X. et al. Convergence of microbial assimilations of soil carbon, nitrogen, phosphorus, and sulfur in terrestrial ecosystems. *Sci. Rep.* **5**, 1–8 (2020).
44. Legendre, P. & Anderson, M. J. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* **69**, 1–24 (1999).
45. Ranjard, L. et al. Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nat. Commun.* **4**, 1–10 (2013).
46. Shanmugam, G. Process-sedimentological challenges in distinguishing paleo-tsunami deposits. *Nat. Hazards* **63**, 5–30 (2012).
47. Szczuciński, W. et al. Sediment sources and sedimentation processes of 2011 Tohoku-oki tsunami deposits on the Sendai Plain, Japan – Insights from diatoms, nannoliths and grain size distribution. *Sediment. Geol.* **282**, 40–56 (2012).
48. Costa, P. J. M. et al. The application of microtextural and heavy mineral analysis to discriminate between storm and tsunami deposits. *Geol. Soc. Spec. Publ.* **456**, 167–190 (2018).
49. Dominey-Howes, D., Dawson, A. & Smith, D. Late Holocene coastal tectonics at Falasarna, western Crete: a sedimentary study. *Geol. Soc. Spec. Publ.* **146**, 343–352 (1999).
50. Switzer, A. D. & Jones, B. G. Large-scale washover sedimentation in a freshwater lagoon from the southeast Australian coast: sea-level change, tsunami or exceptionally large storm? *Holocene* **18**, 787–803 (2008).
51. Waring, B. & Hawkes, C. V. Ecological mechanisms underlying soil bacterial responses to rainfall along a steep natural precipitation gradient. *FEMS Microbiol. Ecol.* **94**, fyy001 (2018).
52. Chénard, C. et al. Temporal and spatial dynamics of Bacteria, Archaea and protists in equatorial coastal waters. *Sci. Rep.* **9**, 1–13 (2019).
53. Saxena, G. et al. Metagenomics reveals the influence of land use and rain on the benthic microbial communities in a tropical urban waterway. *mSystems* **3**, e00136–17 (2018).
54. Hadziavdic, K. et al. Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS ONE* **9**, e87624 (2014).
55. Mariadassou, M., Pichon, S. & Ebert, D. Microbial ecosystems are dominated by specialist taxa. *Ecol. Lett.* **18**, 974–982 (2015).
56. Sheth, A., Sanyal, S., Jaiswal, A. & Gandhi, P. Effects of the December 2004 India Ocean Tsunami on the Indian mainland. *Earthq. Spectra* **22**, S435–S473 (2006).
57. Blot, S. J. & Pye, K. GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surf. Proc. Land.* **26**, 1237–1248 (2001).
58. Folk, R. L. & Ward, W. C. Brazos river bar: a study in the significance of grain size parameter. *J. Sediment. Res.* **27**, 3–26 (1957).
59. Sambrook, J., Russell, D., & Sambrook, J. in *The Condensed Protocols from Molecular Cloning: A Laboratory Manual* (eds Sambrook, J. & Russell, D. W.) (Cold Spring Harbor Laboratory Press, 2006).
60. Wilkins, D., Van Sebille, E., Rintoul, S. R., Lauro, F. M. & Cavicchioli, R. Advection shapes Southern Ocean microbial assemblages independent of distance and environment effects. *Nat. Commun.* **4**, 1–7 (2013).
61. Allen, M. A. & Cavicchioli, R. Microbial communities of aquatic environments on Heard Island characterized by pyrotag sequencing and environmental data. *Sci. Rep.* **7**, 1–16 (2017).
62. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* **17**, 10–12 (2011).
63. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
64. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* **11**, 2639–2643 (2017).
65. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
66. Guillou, L. et al. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604 (2012).
67. R Core Team. R: A language and environment for statistical computing. R <https://www.R-project.org/> (2017).
68. Oksanen, J. et al. *vegan: Community Ecology Package*. Vienna: R Foundation for Statistical Computing. [Google Scholar]. (2016).
69. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26**, 32–46 (2001).
70. Anderson, M. & Ter Braak, C. Permutation tests for multi-factorial analysis of variance. *J. Stat. Comput. Sim.* **73**, 85–113 (2003).
71. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
72. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B: Met.* **57**, 289–300 (1995).
73. Murtagh, F. & Legendre, P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J. Classif.* **31**, 274–295 (2014).

### Acknowledgements

The authors thank F. Constancias and D. Wardle for their insightful discussion in initial manuscript preparation, C. Chénard, E. Acerbi, and P.T. Dat for their valuable advice in initial data processing and analysis. We acknowledge J. Yeo, Dr. S. Krishna Kumar, Dr. R. Krishna Kumar, and Dr. P. Saravanan for helping us collect samples at Cuddalore, India. We would like to send our gratitude to D. Vulot for his constructive comments on the near final versions of this manuscript and his guidance on using DADA2 and phyloseq package in R. We appreciate Dr. S. Chua for wordsmithing the final version of this manuscript. We are grateful to Pedro Costa and two anonymous reviewers; their constructive and insightful reviews significantly improved the manuscript. This research was funded by the Singapore National Research Foundation fellowship scheme (Grant No: NRF-RF2010-04) and the Singapore Ministry of Education, Academic Research Fund Tier 1 RG142/18. This research was also supported by research capacity and infrastructure at Earth Observatory of Singapore and Singapore Centre for Environmental Life Science Engineering, Nanyang Technological University, Singapore that are both funded by National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centre of Excellence initiative.

### Author contributions

W.Y., A.D.S., and F.M.L. conducted the study and wrote the manuscript with contribution from all other co-authors. A.D.S., F.M.L., C.G., D.D., and M.L. conceptualized the idea and contributed to its formality and development. W.Y., A.D.S., and C.G. collected the sediment samples from the field. K.J. provided the access to the sampling site in Thailand. S.S. provided the access to India sampling site and assistance to collect samples in Cuddalore. W.Y. and W.W. conducted molecular laboratory work while W.Y. and Y.Y.T. conducted grain size analysis. W.Y. processed and analyzed the data. E.M. supervised the application and discussion on statistical analysis. W.Y., A.D.S., and F.M.L. contributed to the discussion and interpretation of the results. All co-authors edited and approved the final version of the paper.

### Competing interests

The authors declare no competing interests. A.D.S. is an Editorial Board Member for *Communications Earth & Environment*, but was not involved in the editorial review of, or the decision to publish this article.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-021-00199-3>.

**Correspondence** and requests for materials should be addressed to A.D.S. or F.M.L.

**Peer review information** *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Joe Aslin. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**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) 2021