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

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

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.
he rapid growth of coastal populations significantly increases the exposure of people and assets to coastal hazards such as tsunamis and storm surges\textsuperscript{1,2}. Therefore, assessing coastal hazard risks lies at the core of policy planning and sustainable coastal development\textsuperscript{3,4}. One key piece of information to assess flooding risks in coastal areas is to examine the frequency and magnitude of past events\textsuperscript{1,4,5}. The magnitude and impact of modern events are well captured by instrumental and observational data\textsuperscript{5}. 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\textsuperscript{2,5,6}. 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\textsuperscript{7,8,9,10}.

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\textsuperscript{10–13}. Tsunami and storm surge deposits are frequently indistinguishable due to similarities in sedimentary characteristics, even though they are generated by different mechanisms\textsuperscript{32,14,15}. 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\textsuperscript{9,16}; microfossils such as diatoms, foraminifera, and ostracods may only be preserved for a relatively short time in tropical settings as compared to temperate locations\textsuperscript{17–19}, 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\textsuperscript{20} and microbial activities\textsuperscript{21}, or if the elemental sources are ambiguous\textsuperscript{16}. In addition, compiling multiple proxies might not unequivocally show that deposits were laid down by overwash processes\textsuperscript{22}.

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)\textsuperscript{23} due to catastrophic events such as coastal flooding\textsuperscript{21,24–29}, wildfire\textsuperscript{30,31}, earthquakes\textsuperscript{32,33}. The primary aim of such studies is commonly to investigate how microbes respond to a disturbance in nutrients\textsuperscript{21,28–30}, water resources\textsuperscript{24,25}, substrate availability\textsuperscript{25,27–32}, or other physical environment changes\textsuperscript{23,31–33}.

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\textsuperscript{24–27}. 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\textsuperscript{34,35}. Somboonna et al.\textsuperscript{29} and Asano et al.\textsuperscript{21,28} 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 unflushed terrestrial sediment samples, or marine and intertidal sediment before and after the flooding event (except\textsuperscript{36,37}), 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 barrier on a sandy beach ridge (Fig. 1c). At the India 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\textsuperscript{5,9}, 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\textsuperscript{36}. 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\textsuperscript{5,9,16} and Cuddalore, Tamil Nadu, India\textsuperscript{38–40} (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). In contrast, the Cuddalore system has lower environmental diversity and is dominated by the sandy sediments characteristic of a coastal beach-dune and associated
| Events          | Specific site location (Country, bay/island/prefecture) | Sample descriptions                                                                 | Storage condition | Methodology for microbial community                                      | Key findings                                                                                                                                                                                                 |
|-----------------|---------------------------------------------------------|-------------------------------------------------------------------------------------|------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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.           | • 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. 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. |
| 2004 Indian Ocean Tsunami | Andaman and Nicobar Island, South Andaman: Mithakhari, Crikadabad, Gupatara, 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.           | • Tsunami-affected land had a higher pH in the surface soil layer compared to subsurface as closed to land not affected by tsunami. Soil microbes in unaffected land had a higher culturable microbial population compared to tsunami-affected land. |
| 2004 Indian Ocean Tsunami | India, southeast coast between Vanagiri and Nagoro 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.                          | 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.  Pre-tsunami sediment samples had higher culturable microbial diversity than post-tsunami samples. There were more pathogenic species (bacteria and fungi) in the pre-tsunami sediments. |
| 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 and 18S rRNA using 454 pyrosequencing technique. | 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. 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. The habitat prediction based on percent of species indicators for marine, brackish, freshwater and terrestrial niches indicated that S1 largely comprised of marine-habitat indicator species. Tsunami-affected soil has more strains isolated than the unaffected soil. The isolated strains were found to belong to a single genus, Arthrobacter. 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. 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. |

**Table 1 Summary of all microbial studies on tsunami deposits.**
measure the abundances and evenness of the microbial community using amplicon sequence variants (ASVs). Samples were grouped into a total of 25,034 amplicon sequence variants (ASVs) generated by 16S rRNA barcoding generated an average of 84,011 DNA sequences that were used to generate foraminifera and diatom analyses of foraminifera and diatoms. Metabarcoding foraminifera-specific gene US-pyrosequencing technique.

**Table 1 (continued)**

| Events              | Specific site location          | Sample descriptions                                      | Storage condition | Methodology for microbial community                                      | Key findings                                                                 |
|---------------------|--------------------------------|-----------------------------------------------------------|-------------------|---------------------------------------------------------------------------|----------------------------------------------------------------------------|
| 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 using 454 pyrosequencing technique.                | • Unflooded soil had higher microbial diversity compared to the flooded soil. |
|                     |                                |                                                           |                   |                                                                           | • 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.                                       |
|                     |                                |                                                           |                   |                                                                           | • The effects of the tsunami on soil bacteria in agriculture fields may     |
|                     |                                |                                                           |                   |                                                                           |     have lasted at least 1 year.                                           |
|                     |                                |                                                           |                   |                                                                           | • Sulfur-oxidizing bacteria (SOB), nitrite-oxidizing bacteria (NOB) and    |
|                     |                                |                                                           |                   |                                                                           |     zeta-proteobacteria may serve as an indicator of seawater inundation.   |
| 2011 Tohoku Tsunami | Japan, Hokkaido Island, Urahoro. | A 1.5 m-long sediment core was retrieved and sampled in the field. Seven samples from palaesol samples, 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  | • 172 diatom species were identified with limited fossilized foraminifera    |
|                     |                                |                                                           |                   | foraminifera-specific gene using illumina sequencing technique.           |       tests and radiolaria detected from the samples.                        |
| 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 months and long-term flooded and cultivated field. | −30 °C             | Metabarcoding 16S rRNA using 454 pyrosequencing technique.                | • Foraminifera DNA sequences were detected in all modern samples and most of |
|                     |                                |                                                           |                   |                                                                           |     the palaeo-deposits.                                                     |
|                     |                                |                                                           |                   |                                                                           | • The majority of the foraminifera DNA belonged to brackish and marine taxa. |
|                     |                                |                                                           |                   |                                                                           | • There was limited microbial differences between field flooded for 2       |
|                     |                                |                                                           |                   |                                                                           |     months and unflooded field.                                             |
|                     |                                |                                                           |                   |                                                                           | • There were higher number of sulfur-oxidizing bacteria (SOB) and halotolerant SOB in the field flooded for 2 months. |
|                     |                                |                                                           |                   |                                                                           | • Tsunami flooding created a unique environment that house halotolerant    |
|                     |                                |                                                           |                   |                                                                           |     bacteria that are not found in marine sediments or agricultural soils.  |

*DNA is deoxyribonucleic acid, here refer as genetic material extracted from the environmental samples that were later sequenced to generate the DNA sequences data

DNA is ribosomal ribonucleic acid, it is commonly used in microbiology studies as a tool to characterize unknown microbial organisms

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. 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 Nitrospumilales at both locations, with a much different microbial community structure compared to the higher diversity samples on Phra Thong Island.
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 data
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).

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
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—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 (≤0.02%) and TS (≤0.01%) (Supplementary Fig. 5). In the sedimentary

| Shannon | Simpson |
|---------|---------|
| 0.96    | 0.99    |
| 0.97    | 0.98    |
| 0.98    | 0.99    |
| 0.99    | 1.00    |

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.
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. 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.

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 Nitrosomonadaceae, 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 Micromonosporacae, family Bacillaceae, family Nocardioidaceae, family Sporichthya- ceae, family Caulobacteriaceae and family Sericytchomatiaceae (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
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\(^5\),\(^16\), while Cuddalore’s sediment is mainly sand with limited differences in grain size distribution and characteristically low organic and carbonate content\(^39\) (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\(^45\).

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 metabarcoding 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...
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 < 0.001; 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.25, Asano et al.28, Godson et al.26, and Somboonna et al.29, we observed that tsunami deposits and storm deposits are distinguishable from each other due to differences in the microbial community composition.

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.
deposits have relatively lower diversity as compared to non-
tsunami soils and sediments (Fig. 3). Likewise, Ramesh et al.24
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.27 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 surface50. It could
also affect the incorporation of underlying sediments during
tsunami backflow to the sea4.

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 sta-
tistically 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 mechanisms53
but both extreme waves can transport sediment from the near-
shore marine environment57,48. Some tsunami waves have been
reported to contain offshore sediments from inner contin-
tinal shelf environments59,60 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 communities57,64, either directly (e.g. transporting unique
microbial communities) or indirectly (e.g. altering the soil
chemistry)52,53.

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 eukaryote54. 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 eukaryotes54, while next-generation
sequencing technology has restricted sequencing length. Instead
of using a community-level approach, we suggest a meta-
barcoding 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). Sig-
nature taxa that can identify storm and tsunami deposits live in a
vast range of habitats. This result possibly suggests that micro-
organisms with the capacity to tolerate a broad range of envi-
ronmental conditions increased in relative abundance after the
pulse disturbance of the overwash as opposed to “habitat
specialists”53, 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 commu-
nity changes after the flooding disturbance remained up to 10 years
after the event51.

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 loca-
tions. 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 char-
acterizing 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 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 deposits25,26, the site was used in over 10 studies examining tsunami events27. The 2004 IOT deposits are preserved in the marly swale—Swale X (9° 8.7142` N, 98° 15.916` E) and Swale Y (9° 7.917` N, 98° 15.754` E) in Fig. 1b(i), c, d; Supplementary Data 1), underneath an organic soil layer28. 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, backswab, 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, collecting sediment samples from the geological record, we also collected environmental samples; beach and intertidal sediments using sterile 50 mL conical tube, and carbonate content. Underneath the surface aeolian sediment is the 2011 Cyclone Thane deposits38,40. 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 mineral layers (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), e(ii); Supplementary Data 1), as they contain multiple sources. Sediment samples collected from each sample to prevent cross-contamination between samples. We worn 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, 80° 47.334’ E; Fig. 1b(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 deposits16. Our study site is 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 damage29. Anna University provided permission and field assistance in sample collection from the study site. We pretreated all the tools with 20% bleach solution and rinsed them between each sample to prevent cross-contamination between samples. We used a sterile 50 mL conical tube. For all samples, all tools such as coring tube, Van Veen grab, backswab, hand trowel and Van Dorn bottle were pretreated with 20% bleach solution is used to collect lagoon and marine water samples. The water samples were filtered through a 0.20 μm sterile filters 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 (H2O2). 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 Masterizer 3000. The Malvern Masterizer 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 9.130. This program calculates grain size statistic using the Folk and Ward and moments method31, and generate statistics including the mean(ϕ) 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 each sample was weighed in a pre-weighed tin 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: Chloroformo:isoamyl alcohol (25:24:1)32 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 (ZYMOT Bioresearch, Irvine, CA) to remove any potential impurities. We used the One Step™ DNA Purification Kit to isolate the DNA 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 primer sets: 926F (AAA ATY TAA ATT GGC GTC GG) and 1392R (ACG GGC GGT GTG TRC) is a universal primer that amplifies the 16S rDNA gene V6-V8 regions in archaea, bacteria and eukaryotes33,34; while 574F (GCG GTA ATT CCA GCT CCA A) and 1132R (CCG TAA THT TTV YTT TTT CC) amplifies the 18S rRNA gene V4-V7 regions in eukaryotes54. 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. annealing and polymerization of the primer 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 at the complementary strand at 72°C for 1 min. The following 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 2.2.5. The sequences were then trimmed by truncating those sequences 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^-5 * (quality score)) of higher than 2 (R1) and 5 (R2) using the filterAndTrim function in DADA2 package in R. We applied DADA235 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% similarity36. Taxonomy classification was performed using SILVA version 132 database37 as a reference to identify bacteria and archaea for the 16S rRNA dataset, and PR2 database version 4.12.066 (https://github.com/pr2data-base/pr2database) as a reference to identify eukaryote using RDP naïve Bayesian classification. 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 2926F-1392R primers and 5840 ASVs from 574F-1132R primers.

**Statistical analysis.** Statistical analysis was performed in R version 3.6.1 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 the read length normalized depth. The richness of the samples was calculated using the Shannon index (H’) and Simpson index (D) available in the vegan package38 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 performing statistical analysis. The number of ASVs was transformed with square-root to minimize the effect of abundances.
Research 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 PRJNA343068. 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

1. 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).
2. Gordon, M. et al. in Global Assessment Report on Disaster Risk Reduction Ch 3 (UNDRR, 2019).
3. Domény-Howes, D. Documentory 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).
4. Switzer, A. D., Yu, F., Gouramanis, C., Soria, J. & Pham, T. D. Integrated analysis of tsunami deposits along the Kuril Islands G. & Jaffé, B. E. Physical criteria for distinguishing sandy tsunami and storm deposits using modern samples. Sediment. Geol. 200, 184–207 (2007).
5. Marriner, N. et al. Tsunamis in the geological record: Making waves with a cautionary tale from the Mediterranean. Sci. Adv. 3, e1700485 (2017).
6. Vott, 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 6th cent. AD and a preceding pre-historical event in the early 8th cent. BC. Zeitschrift Geomorphologie
7. Donnelly, J. P. & Woodruff, J. D. Intense hurricane activity over the past 5,000 years controlled by El Nino and the West African monsoon. Nature 447, 465–468 (2007).
8. Tan, N. et al. Impact of fire on active layer and permafrost microbial communities and metagenomes in an upland Alaskan boreal forest. ISME J 8, 195–228 (2014).
9. Moore, C. Y. et al. Tsunamis and storms in coastal areas of the world: observations and inferences. Science 349, 225–264 (2015).
10. Liu, K. B. & Fearn, M. L. Reconstruction of prehistoric landfall frequencies of catastrophic hurricanes in northwestern Florida from lake sediment records. Quat. Res. 54, 238–246 (2000).
11. Morimura, S., Zeng, X., Noboru, N. & Hosono, T. Changes to the microbial communities and metagenomes in an upland Alaskan boreal forest. Annu. Rev. Microbiol. 40, 337–365 (1986).
12. Ramesh, S., Jayaprakashvel, M. & Mathivanan, N. Microbial status in seawater affected terrestrials. Sci. Rep. 7, 198 (2017).
13. 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. J. Hydrol. 468, 460–469 (2014).
14. 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).
15. Pham, D. T. et al. Elemental and mineralogical analysis of marine and coastal sediments from Phra Thong Island, Thailand: Insights into the provenance of prehistoric coastal lagoon ash deposits. Mar. Geol. 385, 274–292 (2017).
16. Shanmugam, G. The tsunami problem. J. Sediment. Res. 76, 718–730 (2006).
17. Curtiss dissimilarity 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 permutation multivariate analyses of variance (PERMANOVA) 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 (Goudaure; appropriate for one-factor analyses) or restricted permutations of residuals (Phra Thong Island; appropriate for multi-factorial analyses). A posteriori pairwise comparisons were done to determine which sample types differ from each other, 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 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) 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.

We applied PCoA calculated using the Bray–Curtis dissimilarity index available in the vegan package to package the similarity and dissimilarity between sample types. We included the result from the Enviﬁt analysis (also available in the vegan package) into the PCoA plot to understand whether the dissimilarity between different sample types is driven by changes in the environment chemistry. The Enviﬁt 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 permutation multivariate analyses of variance (PERMANOVA) 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 (Goudaure; appropriate for one-factor analyses) or restricted permutations of residuals (Phra Thong Island; appropriate for multi-factorial analyses). A posteriori pairwise comparisons were done to determine which sample types differ from each other, 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 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) 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.

Research summary: Further information on research design is available in the Nature Research Reporting Summary linked to this article.
35. Handelmann, J. Metagenomics: application of genomics to uncultured microorganisms. Microbiol. Mol. Biol. Rev. 68, 669–685 (2004).
36. Szczuciński, W. et al. Temporal and spatial dynamics of sedimentary DNA reveals past tsunami deposits. Mar. Geol. 381, 29–33 (2016).
37. Nealson, 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. Conversion 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, nanoliths 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. Dominy-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, fy001 (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, e0136–17 (2018).
54. Hadzic, Z. k et al. Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PLoS ONE 9, e87624 (2014).
55. Maridassou, 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, 2639–2645 (2017).
65. Qust, 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 Protiost Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D742–D746 (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 community data. Ecology 86, 1102–1116 (2005).
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. L., 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. R. 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. Krishnababu, 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. Vaulot 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
G.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.I. 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.W. 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.
