

# SCIENTIFIC REPORTS



OPEN

## Dispersal-competition tradeoff in microbiomes in the quest for land colonization

Francisco Dini-Andreote<sup>1</sup>, Jan Dirk van Elsas<sup>1</sup>, Han Olff<sup>2</sup> & Joana Falcão Salles<sup>1</sup>

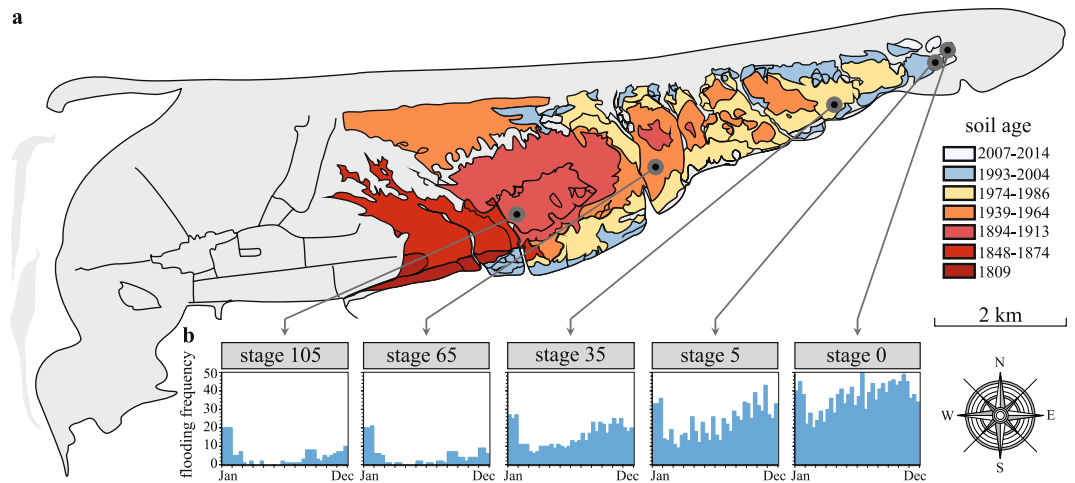
Ancestor microbes started colonizing inland habitats approximately 2.7 to 3.5 billion years ago. With some exceptions, the key physiological adaptations of microbiomes associated with marine-to-land transitions have remained elusive. This is essentially caused by the lack of suitable systems that depict changes in microbiomes across sufficiently large time scales. Here, we investigate the adaptive routes taken by microbiomes along a contemporary gradient of land formation. Using functional trait-based metagenomics, we show that a switch from a microbial 'dispersal' to a 'competition' response *modus* best characterizes the microbial trait changes during this eco-evolutionary trajectory. The 'dispersal' *modus* prevails in microbiomes at the boundary sites between land and sea. It encompasses traits conferring cell chemosensory and motile behaviors, thus allowing the local microbes to exploit short-lived nutritional patches in high-diffusion microhabitats. A systematic transition towards the 'competition' *modus* occurs progressively as the soil matures, which is likely due to forces of viscosity or strain that favor traits for competition and chemical defense. Concomitantly, progressive increases in the abundances of genes encoding antibiotic resistance and complex organic substrate degradation were found. Our findings constitute a novel perspective on the ecology and evolution of microbiome traits, tracking back one of the most seminal transitions in the evolutionary history of life.

The marine-to-terrestrial transition has driven the adaptation of the major forms of life on Earth. In particular, the microbiomes of coastal habitats have been subjected to such transitions as from early evolutionary time<sup>1</sup>, when ecophysiological trait selection allowed microorganisms to thrive under local conditions characterized by progressive increases in the levels of heterogeneity – i.e. changes in microhabitat structure with concurrent increases in oxygenation levels. In such local microhabitats, interactive processes between microbes (e.g. competitive and mutualistic interactions) were presumably on the basis of the emergence of communities with different prevailing traits<sup>2</sup>. However, the eco-evolutionary mechanisms underlying microbiome adaptation to terrestrial life are, thus far, greatly understudied, mainly because of (i) the lack of a fossil record for most microbial groups, and (ii) the absence of studies probing this eco-evolutionary transition along analogous contemporary gradients of soil formation.

From a contemporary ecological perspective, gradients of soil formation (i.e. soil chronosequences), such as those found in coastal ecosystems, are well-suited systems to examine the drivers of trait-based adaptation of local microbiomes across comprehensive time scales. This approach particularly invokes the theme of ecological succession, where communities change over time in a dynamic, and, most importantly, systematic and progressive manner<sup>3,4</sup>. Gradients of naturally forming soil chronosequences can expand over hundreds to thousands of years<sup>5</sup>. In these systems, the temporal scale is based on the assumption of space-for-time substitutions. That is, ecological succession is investigated taking as study objects spatially-distinct microbiomes that share the historical contingency of local habitat genesis (i.e. geogenesis). This fundamental aspect links the ecological basis of this approach to the evolutionary continuous and dynamic trajectory of the soil formation.

A recent surge in the literature has highlighted the importance of understanding the relative contributions of microbial traits to community changes and how these are distributed across divergent community types<sup>6,7</sup>. Notably, functional traits can be measured at various levels of organization – from individual cells, to species, to

<sup>1</sup>Microbial Ecology cluster, Genomics Research in Ecology and Evolution in Nature (GREEN), Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, The Netherlands. <sup>2</sup>Conservation Ecology group, Genomics Research in Ecology and Evolution in Nature (GREEN), Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, The Netherlands. Correspondence and requests for materials should be addressed to F.D.-A. (email: [andreote@gmail.com](mailto:andreote@gmail.com))



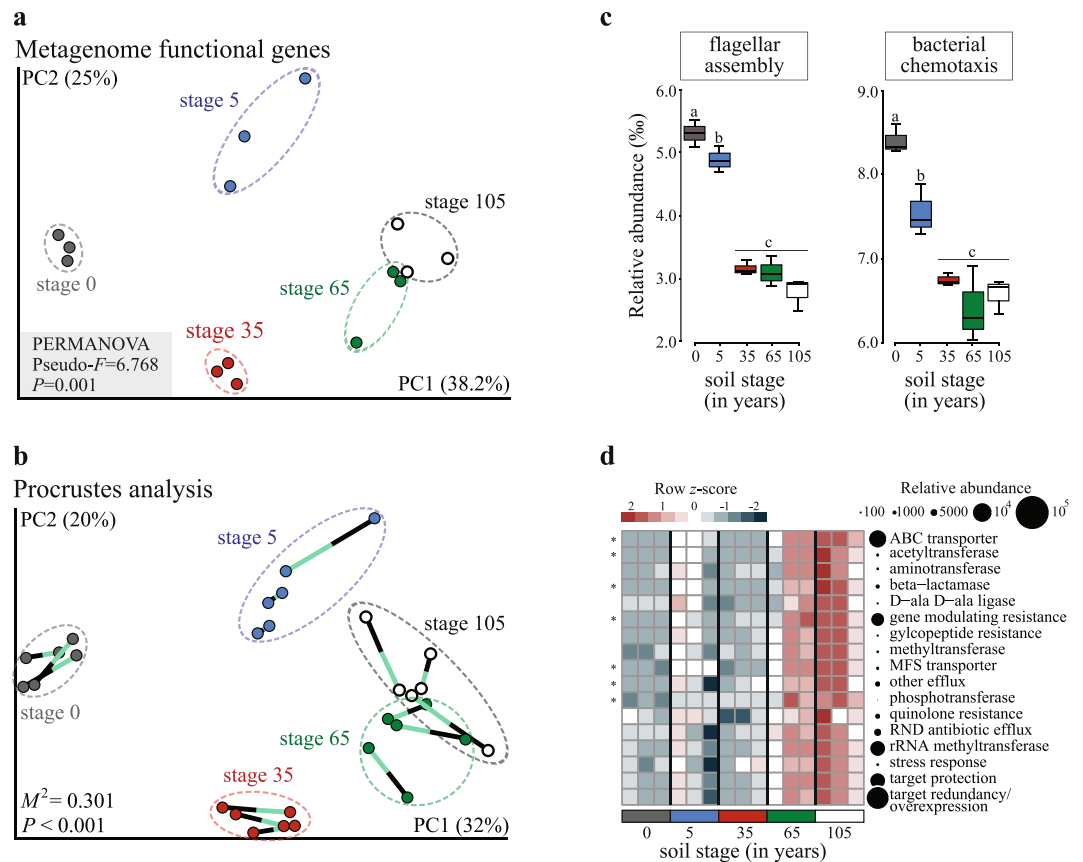
**Figure 1.** Map of the island of Schiermonnikoog, The Netherlands (N53°30' E6°10'). **(a)** The chronology of the geogenesis along the gradual eastward extension of the island, dating from 1809 to 2014 reconstructed from topographic maps, aerial photos and ground-based vegetation maps. Different stages of soil formation were precisely located in the field (see Materials and Methods for details), named as stages 0, 5, 35, 65 and 105 (in years of soil development) (see refs<sup>12–14</sup>). **(b)** Bar charts depict the flooding frequency (i.e., number of times each plot was flooded) per month along the year of 2012 (scaled 0 to 50). The map was created using the software ArcGIS Pro v. 2.0 (<http://pro.arcgis.com/en/pro-app/>).

whole communities<sup>6</sup>. The latter, denoted as ‘community-level traits’, can be quantified as community-aggregated traits (CATs), where traits are measured in a random sample of individuals irrespective of their taxonomic identities. Such a ‘top-down’ approach has proven to yield useful predictors of community-level properties, ranging from soil<sup>8,9</sup> to marine waters<sup>10</sup> to host-associated gut systems<sup>11</sup>.

Here we interrogate whether and how progression in soil formation in an undisturbed marine-to-land gradient leads to shifts in the dominance of prevailing soil microbial community-traits. We thus investigated the microbiomes in a soil formation chronosequence at the island of Schiermonnikoog (The Netherlands), a pristine ecosystem that evolves with a trackable rate<sup>12,13</sup>. The chronosequence covers an area of approximately 8 km long that spans over a century of an ecosystem development<sup>12,13</sup>. The system has been carefully validated with respect to its space-for-time replacement, i.e. all stages of the chronosequence were found to have a similar temporal development<sup>13,14</sup>. In addition, stable isotope analyses ( $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ ) revealed a gradual shift from a system dominated by (external) marine carbon and nutrient inputs to a system dominated by (internal) terrestrial nutrient cycling<sup>15</sup>. In this system, geogenesis occurs through sequential sedimentation of silt and clay particles that are carried by the tides (Figs 1a and S1), resulting in a progressive eastward extension of the island<sup>12</sup>. Through time, the ongoing particle sedimentation causes a progressive decrease of flooding frequency and an increasingly finer soil texture, reducing the diffusion processes in the soil microhabitats (Fig. 1b). Flooding frequency, sodium level and resource availability (particularly dissolved organic carbon) were found to be key determinants of microbial community assembly along this chronosequence, leading to phylogenetically distinct microbiomes across sites<sup>16,17</sup>. Here, we hypothesize that the process of soil formation imposes variable selective pressures on the abundance of specific microbial CATs, leading to progressive shifts in the prevailing ecological strategies in the soil microbiomes. This hypothesis was built on the principle of dispersal-competition trade-off, which constitutes a fundamental ecological synthesis used to explain the coexistence of species in spatially-structured habitats<sup>18,19</sup>, ecological succession dynamics<sup>20</sup>, and abundance patterns of taxa<sup>21</sup>. Dispersal-competition trade-off may also direct the genetic diversity of microbial communities<sup>22</sup>, which often inhabit environments that are seemingly homogeneous but in fact highly structured at the micro-scale. Importantly, this ecological classification reflects similar ideas as the r versus K selection theory of MacArthur and Wilson<sup>23</sup> and the ruderal versus the stress tolerance and competitive strategies of Grime<sup>24</sup>. Specifically, we explored the soil microbiomes with respect to community-level functional traits that match either the ‘dispersal’ (i.e. genes encoding chemotaxis and flagellar motility) or the ‘competition’ *modus* (i.e. genes encoding antibiotic resistance and complex substrate degradation) along an undisturbed gradient of soil formation.

## Results and Discussion

**Soil metagenomes: sequencing and annotation.** To study a putative transition in dominant ecological strategies (that is, examining the ‘dispersal’ versus ‘competition’ *modus*), we used metagenomics data that were produced from the microbiomes sampled along the soil formation chronosequence. We thus extracted total DNA from triplicate soil samples (top 10 cm) along five distinct soil successional stages, i.e. the stages denoted as 0, 5, 35, 65 and 105 years of soil development (Supplementary Table S1). We then performed DNA-based high-throughput sequencing followed by computational sequence analyses. The sequenced metagenomes had sizes of  $1.9 \pm 0.7$  (mean  $\pm$  SD) gigabases (Gb). The raw sequences were quality-trimmed in MG-RAST<sup>25</sup> and annotated using the KEGG Orthology (KO) identifiers<sup>26</sup> ( $E \leq 10^{-5}$ ) (ca.  $12.6 \pm 0.9\%$  of the total reads), and



**Figure 2.** Functional metagenomes correlate with taxonomic composition across the stages of soil formation and the ‘dispersal’ and ‘competition’ response *modus*. **(a)** Principal coordinate analysis (PCoA) based on Bray-Curtis distances calculated from normalized KO annotations. **(b)** Procrustes analysis depicts a significant correlation between functional metagenomes (Bray-Curtis) and bacterial community composition (Bray-Curtis) across the soil stages. **(c)** Box plots display the abundance distribution of flagellar assembly (upper) and chemotaxis (down) mechanisms (based on KO annotations) of microbiomes across the soil stages (asterisks denote independent Mann-Whitney U test results performed only for pairwise comparisons between consecutive soil stages, \* $P<0.05$ , \*\* $P<0.001$ ). **(d)** Heatmap displays the relative abundance (row z-scores) of ARGs by mechanism. ABC, ATP-binding cassette; MFS, major facilitator superfamily; RND, resistance-nodulation-cell division. Circles are proportional to the relative abundance of each ARG mechanism in all samples. Asterisks denote ARG mechanisms that differentially segregated across the soil stages, identified by random forest analysis with Boruta feature selection (average z-scores of 1000 runs  $>4$ ) (Supplementary Table S5).

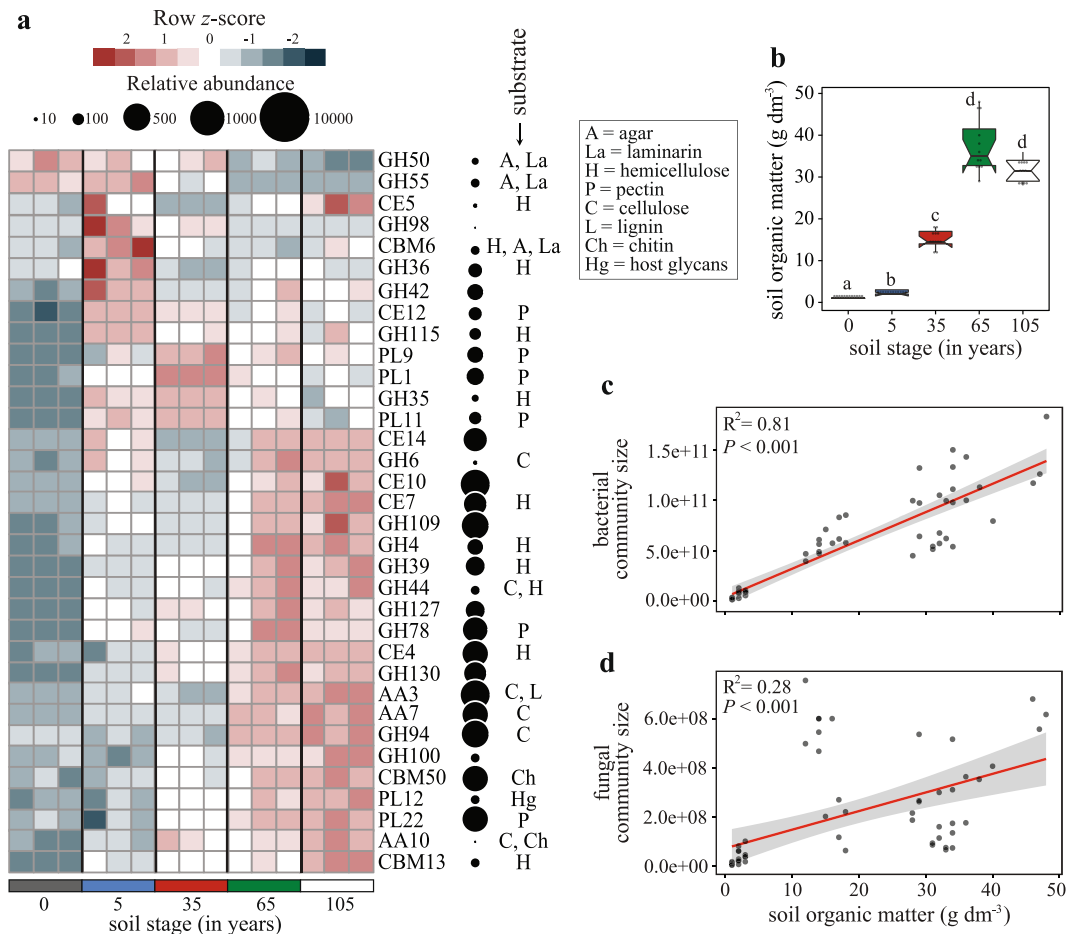
custom-built profile hidden Markov model (HMM) databases for antibiotic resistance genes (ARGs) ( $E \leq 10^{-10}$ )<sup>27</sup> (ca.  $1 \pm 0.1\%$  of the total reads) and for carbohydrate-active enzymes (CAZy) ( $E \leq 10^{-5}$ )<sup>28</sup> (ca.  $0.8 \pm 0.1\%$  of the total reads). Details on the metagenome datasets are provided in Supplementary Table S2.

**Beta-diversity in microbiomes along the gradient of soil formation.** We calculated the differences between the taxonomic (using specific bacterial and fungal markers) and functional (KO, ARG, and CAZy) microbial community structures across the soil chronosequence. The data showed that the taxonomic and functional community structures changed systematically with increasing time of soil formation (Figs 2a and S2a,b, S4a and S5a; Supplementary Table S3), and that these changes were significantly correlated (Fig. 2b;  $P<0.001$ ). Both the functional (KO-based) profiles and the bacterial community compositions (16S rRNA gene-based) clustered according to soil developmental stage, as these consistently displayed highly significant goodness-of-fit measures (RELATE  $\rho=0.782$ ,  $P<0.001$ ; Fig. 2b). Similar results were obtained when the community compositions were compared with the soil resistomes (antibiotic resistance genes [ARG] counts — Supplementary Fig. S4b) and catabolic profiles (CAZy counts — Supplementary Fig. S5b,c and Supplementary Table S4) (Bray-Curtis distances). Collectively, these findings support an evolutionary coupling of phylogeny and function across the different soil microbiomes, and corroborate findings previously reported in the literature<sup>8,27</sup>.

**Evidence for dispersal-competition tradeoff in microbiomes during land colonization.** In line with the ‘dispersal’ *modus*, we suggest that the diffusibility of compounds in the different microhabitats along the soil chronosequence was the key driver of the ability of organisms to either, withstand competition and/or

predation and/or exploit the nutrients or migrate. The flooding frequency in the system decreased systematically along the course of succession (from ca.  $37 \pm 7$  flooding events per month at the early stage, down to  $5 \pm 5$  events at the later one; a full description, including seasonal tidal variations, is provided in Fig. 1b). Thus, we posit that, in the early-stage microhabitats, microorganisms depend on motility as a main life support trait, whereas we expect such traits to lose relevance in habitats in which compound diffusibility is constrained by structure, local viscosity or strain<sup>29</sup>. The *in silico* reconstructions of the genetic pathways for bacterial chemotaxis (PATH: ko02040) and flagellar assembly (PATH: ko02030), per community type across the soil formation gradient, revealed these pathways to be significantly overrepresented at the initial sites (stages 0 and 5), with their relative abundances steadily decreasing towards the late ones (Figs 2c and S3). Thus, in the less-structured soils from the initial stages, where daily tide movements generate constant water fluxes and turnover of high-quality resources (easily decomposable organic matter of marine origin), such traits are thought to assist the local populations in their exploration and exploitation of resource-rich or refuge-like microhabitats. For instance, while in non-turbulent conditions, non-motile cells can explore only ~80 nanoliters of water per day, motile cells can explore over 1 milliliter<sup>30</sup>. This has large implications for the speed at which nutrients are captured in pelagic habitats<sup>30–32</sup>. It is worth mentioning that, in spite of the fact that key groups of marine bacteria (e.g. SAR11 and *Prochlorococcus*) are non-motile, and as such depend exclusively on Brownian motion for movement, bacterial motility is clearly a widespread trait in microbes living in the ocean. Particularly in coastal seawater samples, the occurrence of this trait often ranges from <10 to up to 80% of the total community, varying depending on seasonality, type of system (rich in exploitable surfaces or not), and levels of particulate organic carbon<sup>31,33</sup>. Thus, the transient nutrient patches that occur in these habitats, representing resources that are available in limited space and time, exert a selective force on heterotrophs with respect to motility and chemotaxis traits. Another basis for the overrepresentation of such traits may lie in the quickly changing local conditions, for instance with respect to the levels of toxins, oxygen, redox potential, pH, osmolarity and the intensity and wavelength of light<sup>34,35</sup>. The overrepresentation of ‘dispersal-related’ pathways in the early stages of soil formation therefore likely reflects the intrinsic dynamics of these communities, which requires very quick responses to changing conditions. In these early successional habitats, the short-lived local patchy resources are highly dynamic, e.g. as a result of the filtration of pore water caused by the daily tides. In line with these arguments, our data also revealed a raised number of rRNA operons per organism at the early stages as compared to the later stages of land formation (see ref.<sup>36</sup> for details). These data have been interpreted as a proxy for the ability of bacteria to deal with rapid changes, in particular, increases in nutrient levels. In addition, our findings are consistent with the argument that stochastic processes — leading to temporal niche partitioning — constitute a fundamental ecological mechanism explaining the assembly and turnover of the microbiomes at the early stages of soil formation<sup>16,37</sup>. The community assembly at these early successional sites is probably mainly governed by ecological dispersal. The importance of this process has also been highlighted in other primary and secondary successional gradients. For instance, both in a recently deglaciated soil chronosequence<sup>38,39</sup> and in a post-mining field<sup>40</sup>, ecological dispersal was the main process contributing to the build-up (and restoration) of microbial community diversity at initial (or recently impacted) sites. At these early successional stages, the competition for resource uptake is often still low, and consequently, the effect of competitive exclusion is expected to be minor<sup>41</sup>.

The relative importance of the alternative ecological strategy — the ‘competition’ *modus* — was interrogated by profiling the ability of the microbiomes to resist antibiotics and to degrade complex organic compounds. We hypothesized these traits to be relevant in highly-structured environments, where microorganisms rely — to a greater extent — on chemical warfare for survival. Antibiotics can act as both antimicrobial and signaling molecules, either of these functions benefiting the producer under different conditions<sup>29</sup>. We here focused on the prevalence of resistance to naturally-produced antibiotics. Antibiotic resistance genes (ARGs) occur naturally in microbiomes and may have supported the primordial establishment of bacteria in soils<sup>42</sup>. We examined the prevalence of ARGs of different classes, including predicted drug transporters (e.g. major facilitator superfamily, ABC transporters), antimicrobial modification mechanisms (e.g. acetyltransferases, phosphotransferases) and inactivation mechanisms (e.g. beta-lactamase). The data indicated that, across the soil chronosequence, the ARG patterns had an opposite abundance distribution to that observed for the traits and pathways related to the ‘dispersal’ *modus* (Fig. 2d). This finding is consistent with our dispersal-competition tradeoff hypothesis, in which the increased soil patchiness towards the end of the succession leads to the selection of traits associated with chemical warfare, as opposed to the selection of traits associated with chemosensory motility that are more relevant in the early stages of soil formation. Moreover, the multiplicity of conditions in the soil microhabitats, with different resource levels, is thought to exert significant selective pressures on the local microorganisms with respect to how these deal with competition for local resources. This particular chronosequence is characterized by a progressive increment in soil organic matter (SOM) through time (ca.  $2 \pm 1 \text{ g dm}^{-3}$  at early stages up to  $35 \pm 6 \text{ g dm}^{-3}$  at later ones) (Fig. 3b), resulting from the sediment trapping and organic matter accumulation from the developing vegetation<sup>12</sup>. This increase in SOM was consistent with increments in both the overall bacterial and fungal community sizes (Fig. 3c,d), although more so for the first than for the second group ( $R^2 = 0.81$  and  $R^2 = 0.28$ ,  $P < 0.001$ , respectively). In line with the ‘competition’ *modus*, we thus expected a higher prevalence of genes associated with the degradation of complex compounds in the later stages of soil formation. Moreover, as SOM in the transition zones often consists of a mixture of marine and terrestrial organic inputs<sup>15</sup>, the metabolomes of the local microbiomes were expected to reflect such conditions. Indeed, different categories of carbohydrate metabolism traits were found to dominate along the chronosequence. For instance, traits (as reflected by genes of specific CAZy family proteins) involved in the degradation of labile C sources (e.g., GH50, GH55), often of marine origin in this system, such as agar and laminarin, prevailed in the early stages of the chronosequence. In contrast, traits involved in the degradation of complex plant-derived substrates, including pectin (e.g. PL1, PL9, PL11), and hemicellulose-degrading enzymes (GH35, GH115), dominated at the intermediate stage 35, and those CAZy oxidative enzymes at the late soil stages (e.g., AA3, AA7) (for details see Figs 3a and S5d). Thus, local



**Figure 3.** Soil organic matter (SOM) drives the size of the microbiomes and the distribution of CAZy families through soil formation. **(a)** Heatmap displays the relative abundance (row z-scores) of CAZy gene families that segregated statistically differently across the soil stages, identified by random forest analysis with Boruta feature selection (average z-scores of 1000 runs >4) (Supplementary Table S5). Circles are proportional to the relative abundance of each gene family in all samples. **(b)** Notched box plot depicts the increments in SOM across the soil stages. Different letters indicate significant differences (independent Mann-Whitney U test results performed only for pairwise comparisons between consecutive soil stages,  $P < 0.001$ ). **(c)** Bacterial and **(d)** fungal community sizes (log copy numbers of the maker gene per gram of soil) increments through soil formation. Data are shown as a function of the progressive increments in SOM. The solid lines are linear regression models (contour lines are 95% confidence intervals), and statistics are provided on the panels.

carbohydrate composition was unquestionably a fundamental driver of community functionality and composition. Interestingly, the presence of traits that allow degradation of algal-associated compounds at the early stages was in line with the high abundance of genes associated with bacterial chemotaxis, as such motility systems are known to be stimulated by compounds like those released through algal exudation<sup>33</sup>.

**Methodological considerations.** In the present study, we used a well-validated chronosequence to investigate the systematic changes in microbial CATs along a gradient of soil formation. On the basis of this present-day ecological analog, we provide evidence and theory with respect to how distinct ecological strategies change in importance through the course of soil colonization, thus highlighting shifts in local habitat diffusibility as an important adaptive force in the conquest of terrestrial surfaces by microbial life. Here, we realize that using a present-day chronosequence to explore historical eco-evolutionary transitions in microbiomes is a bold idea that requires a careful interpretation and further exploration. Furthermore, the use of chronosequences assumes that a series of soil sites originated from the same parental material differs solely on the ages since they were formed. This approach does not take into account how changes in environmental conditions and unpredictable disturbances may have interfered in the trajectory of ecological succession. As such, studies on soil chronosequences require a cautious validation of the space-for-time substitution assumption (see Materials and Methods for details). For an additional discussion on the use of chronosequences to investigate long-term ecological succession, see ref.<sup>5</sup>. Finally, the conclusions provided here are built upon relative abundances of selected genes obtained from metagenomics data. Thus, the patterns noted represent the relative investments of the collective microbiome members in the genetic potential for the ‘dispersal’ versus ‘competition’ response *modus*. By definition, ecological traits refer to phenotypic characteristics of organisms, and changes in relative abundances of genes found

in metagenomes do not necessarily equate with changes in CATs. This implies that the use of this approach has inherent limitations to be considered across distinct study types and community functions. For instance, as previously mentioned<sup>6</sup>, genes encoding functions that exhibit a high degree of plasticity should be examined with care, as these may be more prone to be variable in space and time, and across distinct microbiome systems.

**Concluding remarks.** We here provide, for the first time, compelling evidence for the contention that key CATs favoring either a ‘dispersal’ or a ‘competition’ prevailing lifestyle *modus* drive the adaptation of microbiomes along a marine-to-terrestrial transition system. It is important to notice that the fine ecological strategies behind these overall themes (i.e. the type of chemotaxis, antibiotic resistance and/or substrate use behavior) may be diverse at the level of individuals or populations. Future prospective studies may partition such mechanisms and evaluate their quantitative importance across divergent community types, thus providing refined insights into the microbial trait adaptive changes in marine-to-terrestrial transition zones<sup>43</sup>. Moreover, as the composition of the microbiomes differs between each successive soil stage in this system, the differences observed at the level of CATs may have been — to a large extent — caused by shifts in community composition, rather than adaptive changes driven by evolutionary processes. Finally, using a well-validated chronosequence of soil formation to investigate systematic changes in microbiomes is of unique value, as it takes into account the natural formation of the ecosystem, and, as such, accounts its historical contingency. This clearly reflects how biotic and abiotic constraining factors progressively change in the emerging soil gradient and how these changes are reflected in the signatures of CATs in the extant microbiomes. Such an understanding holds a novel, and still unexplored, perspective on how microbiomes are assembled in soils, highlighting a particular set of CATs that underpins their establishment and continuous trajectory in their quest for land colonization.

## Materials and Methods

**Study site.** The black-barrier salt marsh at the island of Schiermonnikoog (N53°30′ E6°10′) represents one of the largest and well-preserved natural salt marshes in Western Europe. From the west to the east, this ecosystem develops along ~8 kilometers encompassing over a century of ecosystem progression<sup>12</sup>. This salt marsh chronosequence has been monitored and calibrated using topographic maps, aerial photographs and the thickness of the sediment layer accumulating on top of the underlying sand layer<sup>12,13</sup>. In addition, permanent plots have monitored the space-for-time replacement of this soil formation at different soil stages for more than 20 years<sup>14</sup>. For this study, five stages were accurately identified and estimated as 0, 5, 35, 65 and 105 years of soil development in 2012 (the sampling year) (referred in the main text and figures as ‘stages 0, 5, 35, 65, 105’). The initial soil sites (stages 0 and 5) are characterized by bare sand and sparsely vegetated soils, respectively. These stages are located in the western part of the chronosequence and are regularly flooded by seawater (Figs 1b and S1). The natural disturbance of the system has been suggested to act as an important mechanism promoting temporal turnover and stochastic assembly of microbial community establishment in these early soils stages<sup>16,17,37</sup>. In particular, the regular flooding acts as a continuous source of immigration of marine-derived microorganisms, in addition to the contribution exerted by the intermittent (initial) input of organic nutrient into the soil. As the ecosystem develops, the chronosequence progresses towards a terrestrial salt marsh by gradually decoupling from the marine environment<sup>15</sup>. Intermediate (stage 35) and late soil sites (stages 65 and 105) are located in elevated areas of the island due to the progressive accretion of mineral particles over time (i.e. silt and clay) — the so-called ‘geogenesis’, and, consequently, these sites are less subjected to the influence of tides. These are well-structured and vegetated soils, having their food-web driven mostly by internal nutrient cycling<sup>15</sup>.

**Soil sampling.** Triplicate plots (5 × 5 m<sup>2</sup>) were established at each identified soil stage (separated 25 m from each other) at the same base of elevation — position at the initial elevation gradient on the bare sand flats with a base elevation of 1.16 m ± 2.2 cm (mean ± SE) above the Dutch ordinance level (NAP). Importantly, differences in the base of elevation reflect differences in inundation regimes, therefore having strong influences on the dynamics and the fate of the ecosystem progression<sup>12</sup>. Soil samples were collected in each plot by randomly taking 20 soil cores (5 cm diameter, 10 cm depth) using sterile techniques, to represent a composite sample. Samples were placed in a sterile plastic bag, sealed and transported to the laboratory (<24 h). All samples were sieved (4 mm mesh size) and stored at –20 °C for total DNA extraction and at 4 °C for physicochemical analysis. Soil samples were collected in May, July, September and November 2012. All samples (*n* = 60) were subject to microbial (bacterial and fungal) community size estimation (quantitative PCR), soil physicochemical characterization and community profiling (bacterial and fungal markers). Shotgun metagenomics was performed for samples collected in July (*n* = 15) 2012.

**Nucleic acid isolation.** Total soil DNA was isolated from 0.5 g of initial material using the MoBio PowerSoil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA). The manufacturer’s protocol was slightly modified by the addition of glass beads (diameter 0.1 mm; 0.25 g) to the soil slurries followed by three cycles of bead beating (mini-bead beater, BioSpec Products, Bartlesville, OK, USA) for 60 s. Obtained DNA samples were quantified using the Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Carlsbad, CA, USA) on a TECAN infinite M200 Pro (Maennedorf, Switzerland) plate reader reading at 485 nm excitation and 530 nm emission. All samples were standardized at equal concentrations for further analysis.

**Soil physical and chemical characterization.** Prior to soil physicochemical characterization, all soil samples were air-dried and sieved through a 100-mesh sieve. Physical (clay:silt:sand %) and chemical determinations of soil organic matter (SOM), nitrate (N-NO<sub>3</sub><sup>-</sup>), ammonium (N-NH<sub>4</sub><sup>+</sup>), sulphate (S-SO<sub>4</sub><sup>2-</sup>), sodium (Na) and pH were performed individually per sample. These determinations were conducted in collaboration with the Laboratory of Soil Analysis at the “Luiz de Queiroz” College of Agriculture (Department of Soil Sciences, ESALQ/USP, Piracicaba, SP, Brazil) according to the methodology described by van Raij *et al.*<sup>44</sup>. Soil N-NH<sub>4</sub><sup>+</sup> and N-NO<sub>3</sub><sup>-</sup> concentrations were determined after extraction in 50 mL of 2 M KCl using 10 g of soil as initial material, according to the methodology described by Keeney and Nelson<sup>45</sup>.

**Flooding frequency estimation.** The frequencies of flooding were calculated by comparing the seawater elevation (measured at the Schiermonnikoog harbor every 10 minutes during the year of 2012) to the natural elevation of each plot. In the Fig. 1b, the y-axis displays the number of times per month each plot (triplicate per site) was estimated to be naturally flooded by sea water (i.e. the number of times that seawater flooded and retracted in each plot — counted as numbers of independent flooding events).

**Shotgun metagenomics.** Shotgun sequencing was conducted following the procedure described in the Illumina TrueSeq DNA sample preparation protocol. Aliquots of each DNA sample were mechanically sheared before entering the library generation. Libraries were size-selected to 170–180 bp using an agarose gel. Sequencing was performed at the Argonne National Laboratory in the Next Generation Sequencing Core (NGS) using a  $2 \times 100$  bp sequencing run on the Illumina HiSeq2000. Raw, unassembled Illumina reads were paired, dereplicated and quality filtered in MG-RAST<sup>25</sup>. Putative open ready frames on the quality-controlled sequences were called using FragGeneScan<sup>46</sup>. Metagenomes were functionally annotated using BLASTX searches against the KEGG (Kyoto Encyclopedia of Genes and Genomes) Orthology (KO) identifiers<sup>26</sup> with a maximum E-value cutoff of  $10^{-5}$ , a minimum percent identity cutoff of 60% and a minimum alignment length cutoff of 15. ARG annotation was performed by searching the amino acid sequence against a previously developed collection of custom-built resistance-gene-specific profile HMMs (<http://dantaslab.wustl.edu/resfams>)<sup>27</sup>, using *hmmsearch* in the HMMER3 software package<sup>47</sup>, with a maximum E-value of  $10^{-10}$ . To generate more encompassing counts of general resistance functions (for example, beta-lactamases), gene counts were summed across all annotations that clearly belonged to the parent function (for example, class A Beta-lactamases, metallo Beta-lactamases, TEM Beta-lactamases), informed by established ARG ontology<sup>48</sup>. CAZy annotation was performed by searching against the HMM profile-based database of carbohydrate-active enzymes obtained from dbCAN v.3<sup>28</sup>, with a maximum E-value of  $10^{-5}$ .

**Statistical analyses.** Cross-soil comparisons were measured using Bray-Curtis similarities calculated based on normalized and square-root transformed count matrices of unique gene sequences per either KOs, ARGs or CAZy families. As the size of the metagenomic libraries varied stochastically by soil sample (Supplementary Table S2), raw counts were normalized to metagenomic library size to account for inconsistent sample depth, facilitating comparison between soils prior to Bray-Curtis calculations. PCoA and PERMANOVA<sup>49</sup> were performed using the homonymous routines in PRIMER6+<sup>50</sup>. Significance levels calculated in PERMANOVA were determined with  $10^3$  permutations. Correlations between resemblance matrices were determined using a non-parametric Mantel-type test implemented as the RELATE routine in PRIMER6+. Procrustes analyses (least-squares orthogonal mapping) were performed in QIIME<sup>51</sup> using two related distance matrices as input. In brief, Procrustes analysis attempts to stretch and rotate the points in one matrix, such as points obtained by principal coordinates analysis (PCoA), to be as close as possible to points in the other matrix, thus preserving the relative distance between points within each matrix<sup>52</sup>. The significance of any Procrustes transformation was determined by comparing the measure of fit,  $M^2$ , between matched-sample PCoA plots to a distribution of  $M^2$  values empirically determined from 10,000 label permutations. In each of the 10,000 permutations, the  $M^2$  value (the sum of squared distances between matched sample pairs) was recalculated and the original  $M^2$  value compared to the simulated distribution in order to compute a *P*-value. Because the  $M^2$  value is dependent on the sample size and data structure, it is generally not comparable across Procrustes transformations. Rather, *P*-values were used to compare different Procrustes plots.

Predicted KOs, ARG mechanisms and CAZy families that segregated significantly between soil stages were identified using random forest analysis<sup>53</sup> with 1,000 trees followed by the Boruta algorithm for feature selection (average *z*-scores of 1000 runs  $>4$ )<sup>54</sup>. These analyses were carried out in R using packages *randomForest* v4.6–7 and *Boruta* v3.0, respectively. Heatmaps were constructed based on *z*-score transformed functional annotations to improve normality and homogeneity of variances. Linear regression modeling and Mann-Whitney U tests were performed using SPSS Statistics version 22 (SPSS, Armonk, NY, USA).

**Bacterial 16S rRNA gene analysis.** Sequencing of the 16S rRNA gene was performed on a Roche 454 GS-FLX automated pyrosequencer running the titanium chemistry as previously described in detail<sup>16</sup>. Briefly, the bacterial 16S rRNA primer set 515f/1061r (targeting the V4–V6 regions) was used given their ability to generate accurate phylogenetic information<sup>55</sup> and having few biases against any bacterial taxa<sup>56</sup>, generating approximately 550-bp reads. All downstream processing was performed in QIIME<sup>51</sup>, for details see ref.<sup>16</sup>.

Quantitative PCR was used to estimate differences in bacterial community size (i.e. as the number of bacterial 16S rRNA gene copies per gram of dry-weight soil) across the soil stages. Quantifications were carried out twice for each soil replicate in 25  $\mu$ L reactions containing 12.5  $\mu$ L Power SYBR Green PCR Master Mix (Applied Biosystems, Frankfurt, Germany), 0.5  $\mu$ L of 20 mg mL<sup>-1</sup> bovine serum albumin (Roche Diagnostics GmbH, Mannheim, Germany), 2  $\mu$ L of each 10 mM primer (FP16S 5'-GGT AGT CYA YGC MST AAA CG-3' and RP16S 5-GAC ARC CAT GCA SCA CCT G-3') targeting the regions V5–V6 of the bacterial 16S rRNA gene<sup>57</sup> and 10 ng of the sample DNA — amplifying a fragment of 263-bp. The thermal cycler protocol was 95 °C for 10 min, 40 cycles of 95 °C for 27 s, 62 °C for 1 min, 72 °C for 30 s, on the ABI Prism 7300 Cycler (Applied Biosystems). The specificity of the amplification products was confirmed by melting curve analysis, and the expected size of the amplified fragment was checked in a 1.5% agarose gel. A standard curve was generated over six orders of magnitude from  $10^3$  to  $10^8$  copies of template per assay ( $R^2 = 0.99$ ) using a plasmid containing the cloned 16S rRNA gene from *Serratia plymuthica*. The quantitative PCR efficiency (*E*) was calculated according to the equation  $E = [10^{(-1/\text{slope})} - 1]$ . Possible inhibitory effects of co-extracted humic compounds were checked by spiking samples with a range of known concentrations of the plasmid. No apparent inhibition was observed.

**Fungal ITS analysis.** Sequencing of the ITS region of the fungal rRNA gene was performed on an Illumina MiSeq paired-end (2 × 250 bp) platform. We used the primer set ITS1f<sup>58</sup> and ITS2<sup>59</sup>, slightly modified for sequencing using the Illumina MiSeq (for a detailed description see ref.<sup>60</sup>). All downstream processing was performed using QIIME<sup>51</sup> and USEARCH v7<sup>61</sup>. The full pipeline is available at, <http://brmicrobiome.org> (see ref.<sup>17</sup>).

Quantifications of fungal community sizes (i.e. as the number of fungal ITS copies per gram of dry-weight soil) were carried out twice for each soil sample in 25 µl reactions containing 12.5 µl Power SYBR Green PCR Master Mix (Applied Biosystems, Frankfurt, Germany), 1.25 µl of 20 mg mL<sup>-1</sup> bovine serum albumin (Roche Diagnostics GmbH, Mannheim, Germany), 1.25 µl of each 10 µM primer (ITS1f 5'-TCC GTA GGT GAA CCT GCG G-3' and 5.8S 5'-CGC TGC GTT CTT CAT CG-3')<sup>58,62</sup> and 10 ng of the sample DNA — amplifying a fragment of ~300-bp. The thermal cycler protocol was 95 °C for 10 min, 40 cycles of 95 °C for 1 min, 53 °C for 30 s, 72 °C for 1 min, on the ABI Prism 7300 Cycler (Applied Biosystems). The specificity of the amplification, the presence of inhibitory compounds and the qPCR efficiency were tested as described for bacteria (see above). The standard curve (R<sup>2</sup> = 0.97) was generated using a plasmid containing the cloned ITS region from *Rhizoctonia solani*. No apparent inhibition was observed.

**Data availability.** The metagenome data sets supporting the results in this article are publically available at the MG-RAST server with the accession numbers 4558897.3, 4558900.3, 4558903.3, 4558906.3, 4558909.3, 4558912.3, 4558915.3, 4558918.3, 4558921.3, 4558924.3, 4558927.3, 4558930.3, 4558933.3, 4558936.3, 4558939.3. The amplicon sequence data sets are available at the EBI ENA with accession numbers PRJEB22178 (ERP024520): bacterial 16S rRNA sequencing data; and PRJEB22169 (ERP024509): fungal ITS sequencing data.

## References

- Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4**, 44 (2004).
- Wu, H., Fang, Y., Yu, J. & Zhang, Z. The quest for a unified view of bacterial land colonization. *ISME J.* **8**, 1358–1369 (2013).
- Clements, F. E. *Plant succession: an analysis of the development of vegetation*. Carnegie Institution of Washington (Washington, DC, USA 1916).
- Walker, L. R. & del Moral, R. *Primary succession and ecosystem rehabilitation*. Cambridge University Press (Cambridge, UK 2003).
- Walker, L. R., Wardle, D. A., Bardgett, R. D. & Clarkson, B. D. The use of chronosequences in studies of ecological succession and soil development. *J. Ecol.* **98**, 725–736 (2010).
- Fierer, N., Barberán, A. & Laughlin, D. C. Seeing the forest for the genes: using metagenomics to infer the aggregated traits of microbial communities. *Front. Microbiol.* **5**, 614 (2014).
- Martiny, J. B., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes in light of traits: a phylogenetic perspective. *Science* **350**, aa9323 (2015).
- Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. USA* **109**, 21390–21395 (2012).
- Dini-Andreote, F., Brossi, M. J., van Elsas, J. D. & Salles, J. F. Reconstructing the genetic potential of the microbially-mediated nitrogen cycle in a salt marsh ecosystem. *Front. Microbiol.* **7**, 902 (2016).
- Ganesh, S., Parris, D. J., Delong, E. F. & Stewart, F. J. Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.* **8**, 187–211 (2014).
- Greenblum, S., Turnbaugh, P. J. & Borenstein, E. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc. Natl. Acad. Sci. USA* **109**, 594–599 (2012).
- Olf, H., De Leeuw, J., Bakker, J. P., Platerink, R. J. & van Wijnen, H. J. Vegetation succession and herbivory in a salt marsh: changes induced by sea level rise and silt deposition along an elevational gradient. *J. Ecol.* **85**, 799–814 (1997).
- Schrama, M., Berg, M. P. & Olf, H. Ecosystem assembly rules: the interplay of green and brown webs during salt marsh succession. *Ecology* **93**, 2353–2364 (2012).
- Van Wijnen, H. J., Bakker, J. P. & de Vries, Y. Twenty years of salt marsh succession on a Dutch coastal barrier island. *J. Coast. Conservat.* **3**, 9–18 (1997).
- Schrama, M., Jouta, J., Berg, M. P. & Olf, H. Food web assembly at the landscape scale: using stable isotopes to reveal changes in trophic structure during succession. *Ecosystems* **16**, 627–638 (2013).
- Dini-Andreote, F. *et al.* Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. *ISME J.* **8**, 1989–2001 (2014).
- Dini-Andreote, F., Pylro, V. S., Baldrian, P., van Elsas, J. D. & Salles, J. F. Ecological succession reveals potential signatures of marine–terrestrial transition in salt marsh fungal communities. *ISME J.* **10**, 1984–1997 (2016).
- Levins, R. & Culver, D. Regional coexistence of species and competition between rare species. *Proc. Natl. Acad. Sci. USA* **68**, 1246–1248 (1971).
- Tilman, D. Competition and biodiversity in spatially structured habitats. *Ecology* **75**, 2–16 (1994).
- Tilman, D. Constraints and tradeoffs: toward a predictive theory of competition and succession. *Oikos* **58**, 3–15 (1990).
- Guo, Q., Brown, J. H., Valone, T. J. & Kachman, S. D. Constraints of seed size on plant distribution and abundance. *Ecology* **81**, 2149–2155 (2000).
- Yawata, Y. *et al.* Competition–dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proc. Natl. Acad. Sci. USA* **111**, 5622–5627 (2014).
- Pianka, E. R. On r and K selection. *Amer. Nat.* **104**, 592–597 (1970).
- Grime, J. P. *et al.* Integrated screening validates primary axes of specialisation in plants. *Oikos* **79**, 259–281 (1997).
- Meyer, F. *et al.* The metagenomics RAST server – a public resource for the phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**, 386 (2008).
- Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**(Database issue), D480–D484 (2008).
- Forsberg, K. J. *et al.* Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**, 612–616 (2014).
- Yin, Y. *et al.* dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
- Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. Bacterial competition: surviving and thriving in the microbial jungle. *Nature Rev. Microbiol.* **8**, 15–25 (2010).
- Stocker, R. Marine microbes see a sea of gradients. *Science* **338**, 628–633 (2012).
- Fenchel, T. Microbial behavior in a heterogeneous world. *Science* **296**, 1068–1071 (2002).
- Blackburn, N., Fenchel, T. & Mitchell, J. Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria. *Science* **282**, 2254–2256 (1998).
- Stocker, R. & Seymour, J. R. Ecology and physics of bacterial chemotaxis in the ocean. *Microbiol. Mol. Biol. Rev.* **76**, 792–812 (2012).
- Wadhams, G. H. & Armitage, J. P. Making sense of it all: bacterial chemotaxis. *Nature Rev. Mol. Cell. Biol.* **5**, 1024–1037 (2004).



35. Dennis, P. G., Seymour, J., Kumbun, K. & Tyson, G. W. Diverse populations of lake water bacteria exhibit chemotaxis towards inorganic nutrients. *ISME J.* **7**, 1661–1664 (2013).
36. Nemergut, D. R. *et al.* Decreases in average bacterial community rRNA operon copy number during succession. *ISME J.* **10**, 1147–1156 (2016).
37. Dini-Andreote, F., Stegen, J. C., van Elsas, J. D. & Salles, J. F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. USA* **112**, E1326–E1332 (2015).
38. Nemergut, D. R. *et al.* Microbial community succession in an unvegetated, recently deglaciated soil. *Microb. Ecol.* **53**, 110–122 (2007).
39. Schmidt, S. K., Nemergut, D. R., Darcy, J. L. & Lynch, R. Do bacterial and fungal communities assemble differently during primary succession? *Mol. Ecol.* **23**, 254–258 (2014).
40. Harantová, L. *et al.* Development of microbial community during primary succession in areas degraded by mining activities. *Land Degrad. Develop.* **28**, 2574–2584 (2017).
41. Lindström, E. S. & Langenheder, S. Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* **4**, 1–9 (2012).
42. D'Costa, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477**, 458–461 (2011).
43. Tout, J. *et al.* Chemotaxis by natural populations of coral reef bacteria. *ISME J.* **9**, 1764–1777 (2015).
44. Van Raij, B., Andrade, J. C., Cantarella, H. & Quaggio, J. A. *Análise química para avaliação da fertilidade de solos tropicais* (Instituto Agronômico de Campinas 2001).
45. Keeney, D. R. & Nelson, D. W. In: *Methods of soil analysis: part 2* (eds Page, A. L., Miller, R. H. & Keeney, D. R.) 643–698 (ASA and SSSA, 1982).
46. Rho, M., Tang, H. & Ye, Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* **38**, e191 (2010).
47. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
48. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–3357 (2013).
49. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46 (2001).
50. Clarke, K. R. & Gorley, R. N. *PRIMER v6: User Manual/Tutorial, 6th Edition, Ch. 13.* (PRIMER-E, 2006).
51. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
52. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**, 970–974 (2011).
53. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
54. Kursa, M. B. & Rudnicki, W. R. Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010).
55. Liu, Z., Lozupone, C., Hamady, M., Bushman, F. D. & Knight, R. Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* **35**, e120 (2007).
56. Bates, S. T. *et al.* Examining the global distribution of dominant archaeal populations in soil. *ISME J.* **5**, 908–917 (2011).
57. Bach, H. J., Tomanova, J., Schloter, M. & Munch, J. C. Enumeration of total bacteria and bacteria with genes for proteolytic activity in pure cultures and in environmental samples by quantitative PCR mediated amplification. *J. Microbiol. Methods* **49**, 235–245 (2002).
58. Gardes, M. & Bruns, T. D. ITS primers with enhanced specificity for basidiomycetes —application to the identification of mycorrhizae and rusts. *Mol. Ecol.* **2**, 113–118 (1993).
59. White, T. J., Bruns, T. D., Lee, S. & Taylor, J. In: *PCR protocols: a guide to method and applications* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 315–322 (San Diego Academic Press, 1990).
60. Smith, D. P. & Peay, K. G. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *Plos One* **9**, e90234 (2014).
61. Edgar, R. C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998 (2013).
62. Vilgalys, R. & Hester, M. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. *J. Bacteriol.* **172**, 4238–4246 (1990).

## Acknowledgements

We thank the “Nederlandse Vereniging voor Natuurmonumenten” for granting us access to the salt marsh area at the island of Schiermonnikoog, The Netherlands. We also thank Diego Chaves for technical assistance in ARG and CAZy annotations. This research was supported by the Netherlands Organization for Scientific Research (NWO) and the Gratama Foundation.

## Author Contributions

F.D.-A., J.D.v.E. and J.F.S. designed the research; J.D.v.E. and J.F.S. contributed with new reagents and analytical tools; F.D.-A. collected the soil samples and performed the research; F.D.-A. analyzed the data with the supervision of J.D.v.E. and J.F.S.; F.D.-A., J.D.v.E., H.O. and J.F.S. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-27783-6>.

**Competing Interests:** The authors declare no competing interests.

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



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

© The Author(s) 2018