Temporal dynamics and ecophysiology of thermophilic composting analyzed through metagenome-assembled genomes

CURRENT STATUS: UNDER REVIEW

João Carlos Setubal
Universidade de Sao Paulo
setubal@iq.usp.br
ORCiD: https://orcid.org/0000-0001-9174-2816

Lucas Palma Perez Braga
Universidade de Sao Paulo

Roberta Verciano Pereira
Universidade de Sao Paulo

Layla Farage Martins
Universidade de Sao Paulo

Livia Maria Silva Moura
Universidade de Sao Paulo

Fabio Beltrame Sanchez
Universidade de Sao Paulo

José Salvatore Leister Patané
Instituto Butantan

Aline Maria da Silva
Universidade de Sao Paulo

DOI:
10.21203/rs.3.rs-19445/v1

SUBJECT AREAS
General Microbiology

KEYWORDS
thermophilic composting, MAGs, metagenome-assembled genomes, Sao Paulo Zoological Park, Brazil
Abstract

Background: Thermophilic composting is a semi-engineered process carried out by diverse microbial communities. Composting is an environment friendly way of degrading biomass; its study may help uncover important biomass-degrading organisms and key enzymes. DNA sequence-based previous studies have presented a general description of the microbial-molecular features of composting, but they have lacked more specific information on the key organisms that are active during the process and their genomes.

Methods: We present an analysis of metagenome-assembled genomes (MAGs) obtained from time-series samples of a thermophilic composting process in the São Paulo Zoological Park (Brazil). Our results are based on a careful analysis of MAG gene content and on metabolic modeling of their interactions.

Results: We recovered 60 MAGs from sequencing datasets from two separate composting cells. Phylogenetic analysis shows that 47 of these MAGs represent novel taxa at the genus or higher levels. We have analyzed the gene repertoire of these MAGs in terms of lignocellulose degradation, secondary metabolite production, antibiotic resistance genes, denitrification genes, sulfur metabolism, hydrogen metabolism, and oxygen metabolism. For one of the composting cells we also had metatranscriptome data, which allowed a deeper analysis of 49 MAGs. This analysis showed the presence of three distinct clusters of MAGs with varying activity during the 99-day composting process. The interaction model pointed to Sphaerobacter thermophilus and Thermobispora bispora as key players in the process, as well as other bacteria that are novel. Our results also show the importance of coadjuvant bacteria and of microbial functions related to efficient bioenergetic processes during biomass conversion, such as N2O reduction and hydrogenases. A novel acidobacteria MAG encodes N2O reductase hallmark genes (nosZD). Strong metabolic dependencies predicted between MAGs revealed that cross-feeding in composting can be determined by complementary functions found in the genomes of producers and consumers, supporting the Black Queen hypothesis for co-evolutionary interactions.

Conclusions: This study reveals for the first time the key bacterial players in thermophilic composting
and provides a model of their dynamic metabolic interactions. These findings pave the way for more rational composting procedures and provide information that could help the development of novel biomass-degrading technologies.

Introduction

Thermophilic composting is a semi-engineered process carried out by microbial communities able to thrive in those environments [1-3]. Composting microbes present a remarkable metabolic flexibility and are able to break down complex compounds such as lignocellulosic biomass [3, 4]. Therefore, the composting microbiome is a valuable microbial resource with potential for contributing in a number of industrial applications [5, 6]. In spite of all this potential, knowledge on how to control and explore those microbes and their functions remains encrypted within their genomes and the multiple combinations of metabolic pathways that they can activate [7-9].

So far, research on composting microbes has focused mainly on enriched cultures [9-11] or taxonomic biodiversity assessments based on 16S rRNA gene amplicon sequencing data [12, 13]. Enriched cultures have lower diversity richness due to cultivation bias and the limited number ecological niches [9]. Community profiling based on 16S rRNA gene amplicon data may reveal insights about microbial composition and the dynamics of ecological succession [14], however those methods cannot assess microbial functional diversity and metabolic activity.

Lignocellulosic biomass is composed by different biopolymers: cellulose (25–55%), hemicellulose (19–40%) and lignin (18–35%), and smaller fractions of pectin and minerals [15]. Therefore, a diverse set of enzymes is required for effective saccharification (i.e., depolymerization of lignocellulose compounds into monosaccharide components) [16]. Microbial dynamics during lignocellulose breakdown seems to be heavily dependent on syntrophic interactions [9]. The microbial populations need to share the burden of enzymatic production; and sharing metabolites reduces the negative feedback effect of intermediate metabolite accumulation (e.g., intracellular enzymatic competition) [17]. Syntrophic interactions happen in the presence of opportunistic microbes in biomass degrading systems; these are microbes that do not express or very often lack the required enzymes for biomass degradation, but constitute one important portion in microbial consortia obtained from enrichment
cultures (sometimes referred to as ‘sugar cheaters’) [9, 18].

Shotgun metagenomic sequencing has helped reveal the diversity of microbial communities in environmental samples [19-21]. New methods and computational tools allow the recovery of metagenome-assembled genomes (MAGs) from complex environments [22-26]. The availability of MAGs from organisms in a given environment provides detailed taxonomical and functional diversity information, and therefore has the potential to allow a better understanding of the ecological context and the arsenal of enzymes encoded by microbes living in that environment [25, 27, 28].

In order to control and fully explore the composting system as a microbial source for biotechnological solutions, it is important to consider the ecological basis of the composting microbiome. Therefore, here we present an analysis of MAGs obtained from time-series samples of a thermophilic composting process [3]. Our aim was obtain a more detailed view of the microbial populations active in a composting process, advancing on our previous work [3].

We used the collection of genomes recovered to build a framework for the molecular-microbial temporal dynamic in the composting process, from which we were able to infer the major ecophysiological patterns of its microbiome. Using this framework as a reference, we built genome-scale metabolic models for predicting syntrophic interactions and the more frequently exchanged compounds.

For ease of reading, in the text we sometimes we refer to MAGs when we actually mean the bacteria from which these genomes were obtained.

Results
MAGs recovered from composting

A total of 11 MAGs from the ZC3 dataset and 49 from the ZC4 dataset were recovered. (Supplementary Table S1). All of our MAGs meet the medium-quality requirement (> 50% completeness and < 10% contamination) of the MIMAG standard [29] (with one exception: ZC4RG07 had 10.87% contamination); 34 MAGs meet the high-quality requirement (> 90% completeness and < 5% contamination). The average number of contigs in these MAGs is 363.75, with a minimum of 15 and a maximum of 1,655. Thirteen MAGs could be assigned to species for which there is at least one
genome publicly available; this allowed us to compare our MAGs to publicly available genomes of the same species (Supplementary Table S2a). In all cases the two-way ANI measure was above 98% and the GGDC DDH estimate (formula 2) was at least 87.3%, strongly suggesting that the assignments are correct and that the recovered genomes are of high quality. These 13 MAGs correspond to 11 species that have been found in environments related to biomass degradation, with the possible exception of M. hassiacum, which has been isolated in human patients. All these species have been reported as thermophilic bacteria (Supplementary Table S3).

**Taxonomic assignments**

The 60 recovered MAGs could be assigned to six different phyla: Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Firmicutes and Proteobacteria (Supplementary Table S1). At the order level there is remarkable diversity: 32 different orders are represented. The most frequent order was Limnochordales (six MAGs). Most of the MAGs seem to be novel: there are seven MAGs for which no order could be assigned, 14 MAGs for which no family could be assigned, 14 MAGs for which no genus could be assigned, and 12 MAGs for which no species could be assigned. This counting does not take into account GTDB-tk assignments to taxa that are not currently accepted (such as family WCHB1–69 within the order bacteroidales).

**Presence of MAGs in other datasets**

We compared the 60 MAGs with genomes recovered from our previous studies [3, 11] and between both composting cells (Supplementary Table S2b). We observed that a few genomes can be said to be present (using the same methodology for MAG species assignment) in different samples, although the majority was found in only one sample. The recovery of the “same” genome from different samples lends additional confidence to our genome recovery process. It is probable that there may be more occurrences of the same MAGs in these samples than here reported, but our sequencing coverage and stringent genome recovery criteria may have prevented us from recovering additional already-observed MAGs from all samples.

**Functional analysis of MAGs**

For functional analysis, we have focused on ZC4 samples only, since for them alone do we have metatranscriptome data. ZC4 is composed of nine time-series samples (days 1, 3, 7, 15, 30, 64, 67,
A comparison of variation in relative abundance of transcripts over time showed that all MAGs here analyzed were transcriptionally active (Supplementary Table S4)

The composting process is clearly very complex, both in terms of its microbiota and in terms of the varied subprocesses that occur over the approximately three months during which composting takes place. For the functional analysis that follows we have focused on functions that we deemed relevant for composting microbial systems. Those that we chose to study in detail are: lignocellulose degradation; denitrification; sulfur metabolism; hydrogen metabolism; and oxygen metabolism.

Secondary metabolite production and antibiotic resistance genes were also included given their role in microbial interactions.

Lignocellulose degradation

Biomass degrading capabilities in MAGs were analyzed based on COG assignment (Supplementary Fig. S1) and CAZy annotation (Fig. 1; and Supplementary Tables S5 and S6). Among the 49 ZC4 MAGs, 12 present each more than 150 CDSs classified as CAZymes (Fig. 1). In these 12 MAGs, each has at least 40 CDSs also classified as GHs (Fig. 1). Several cellulases (GH5, GH6, GH9 and GH45), endohemicellulases (GH8, GH10, GH11, GH12, GH26, GH28 and GH53), debranching (GH51, GH62, GH67 and GH78) and oligosaccharide-degrading enzymes (GH1, GH2, GH3, GH29, GH35, GH38, GH39, GH42 and GH43) were found in these MAGs (Fig. 1).

Regarding Auxiliary Activities (AA), ZC4RG20 (c__Gammaproteobacteria), ZC4RG33 (g__Aquamicrobium), ZC4RG43 (s__Mycobacterium_hassicum), and ZC4RG45 (s__Thermocrispum agresta) present at least 15 CDSs classified as AA (Supplementary Table S6). ZC4RG45 contains the highest diversity of AA genes. Members of the AA1 family, which perform lignin degradation efficiently, were only found in ZC4RG08 (s__Pseudomonas themotolerans). ZC4RG21 (s__Thermobifida fusca), ZC4RG04 (s__Thermobispora bispora), ZC4RG28 (f__Streptosporangiaceae), ZC4RG45, and ZC4RG47 (g__Micromonospora) were the only ones containing CDSs classified in the AA10 family (lytic polysaccharide monooxygenases), members of which are capable of directly targeting cellulose for oxidative cleavage of the glucose chains.

As stated above, we have strong evidence that each MAG here analyzed was transcriptionally active.
We checked the expression of CDSs related to lignocellulose degradation, and determined that all CAZymes mentioned here are being expressed in the composting process (Fig. 2).

**Secondary metabolites**

Several CDSs classified as secondary metabolite genes (siderophores, bacteriocins, sacpeptides, betalactones, lassopeptides, and type I, II and III polyketides) were found in MAGs (Fig. 3). MAGs with at least six secondary metabolites genes were ZC4RG43 (s__M. hassiacum) (16 genes), ZC4RG39 (f__Steroidobacteraceae) (13), ZC4RG21 (s__T. fusca) (10), ZC4RG47 (g__Micromonospora) (7), ZC4RG22 (o__Luteitaleales) (6), ZC4RG04 (s__T. bispora) (6) e ZC4RG46 (o__Polyangiales) (6). Most of these MAGs were classified as Actinobacteria or Proteobacteria.

We observed that secondary metabolite genes were more expressed in the initial days (D1, D3, D7) than the final days (D78, D99) (Supplementary Table S7).

**Antibiotic resistance genes**

Antibiotic resistance gene (ARG) clusters were observed mainly in MAGs from Actinobacteria and Proteobacteria phyla (Fig. 4). ZC4RG08 (s__P. thermotolerans) has the largest number of ARGs (12 CDSs), followed by ZC4RG43 (s__M. hassiacum) (11 CDSs). Additionally, several multidrug efflux pumps were found in ZC4RG08 (MuxABC-OpmB, MexAB-OprM, MexEF-OprN, MexWV and MexJK) (Supplementary Table S8).

Transcripts coding for resistance genes were more abundant in the early (D01 and D03) and final (D78 and D99) days (Supplementary Table S9).

**Aerobic and Anaerobic respiration strategies**

The analysis of oxygen metabolism indicates that nearly all bacteria from which MAGs were obtained are aerobes (Supplementary Table S10). The oxidases detected were all active and transcript abundance variation over time shows a slight decrease in D7, with an increase following the turning procedure (Supplementary Fig. S2). Evidence of oxidases and aerobic metabolism was not detected in MAGs ZC4RG12 (g__Caldicoprobacter), ZC4RG32 (s__Caldicoprobacter faecalis), ZC4RG34 (f__Thermovenabulaceae), and ZC4RG49 (s__[Clostridium] cellulosi), indicating thereby a metabolism strictly anaerobic. These MAGs have all been classified within the phylum Firmicutes and demonstrated a similar profile of activity based on the global abundance variation of their transcripts.
across composting, with a peak in D7 (Supplementary Table S4).

Sulfate reduction via cysteine desulphurase (cysCN) was detected as a widespread and active function among MAGs (Supplementary Table S10). Such mechanism of sulfate reduction is part of the assimilatory sulfate reduction pathway by which sulfate is incorporated into cysteine. Evidence for dissimilatory sulfite reductase function (dsrAB and apsA) was not detected in the genomes, which is evidence that the respiratory sulfate reduction was not the main strategy for anaerobic respiration employed by the bacteria in this composting process.

Active denitrification genes were detected in several MAGs (Supplementary Table S10). Eighteen of them presented the nitrite respiration gene nirK. The variant nirS was not detected in the MAGs. ZC4RG13 (s__Rhodothermus marinus), ZC4RG22 (o__Luteitaleales), ZC4RG26 (s__Sphaerobacter thermophilus), and ZC4RG29 (f__Cyclobacteriaceae) encode the complete denitrification pathway (i.e., nitrous-oxide reductase pathway, nosDZ). ZC4RG29 lacks the nirK gene and has evidence for nitrite reductase using the nirB gene. Nitrous-oxide reductase genes (nosZD) were also active during the composting process. The variation in abundance of transcripts of these denitrification genes increases over time, with a peak starting in D7 (Supplementary Fig. S2).

Chemolithotrophic metabolism
We found evidence for chemolithotrophic metabolism based on MAG genes related to the oxidation of inorganic sulfur compounds (Supplementary Table S10). Nearly all MAGs have genes from the sulfur oxidation pathway via sulfur dioxygenase. Some of the MAGs were found to be more versatile and had genes annotated with other functions associated with the oxidation of sulfur compounds. ZC4RG20 (c__Gammaproteobacteria), for instance, represents a bacterial population that showed transcripts associated with sulfur dioxygenase, sulfide oxidation (sqr), and thiosulfate oxidation (soxC), including transcripts associated with carbon fixation via rubisco activation, supporting a chemolithoautotrophic growth. The Sox system, which is able to oxidize sulfite and sulfone group in thiosulfate, was found in MAGs ZC4RG25 (f__Hyphomicrobiaceae), ZC4RG31 (f__Hyphomicrobiaceae), ZC4RG33 (g__Aquamicrobium), and ZC4RG42 (o__Betaproteobacteria), although soxC was apparently lacking in all of them. Sulfide oxidation (sqr) to thiosulfate was detected also in ZC4RG25. These observations
highlight that members of the bacterial populations in the composting microbiome were capable of harvesting energy by oxidizing inorganic sulfur compounds. CDSs associated with nitrification genes (amo and hao) were not detected in the MAGs, indicating that this trait was of minor or no relevance for the composting microbiome.

Several hydrogenases were found to be present and expressed (Supplementary Table S10). We were able to identify two types of hydrogenases. MAGs ZC4RG04, 09, 13, 15, 26, 28, 36, 37, 38, 43, and 49, belonging to diverse phyla (Supplementary Table S1), presented CDSs associated with [NiFe] hydrogenases. MAGs ZC4RG11, 12, 23, 32, 34, 38, and 49, from the phylum Firmicutes (Supplementary Table S1), have CDSs annotated as prototypical hydrogen-evolving [FeFe] hydrogenases (group A1).

Co-occurrence of MAGs
Using the 49 ZC4 MAGs, we inferred correlation patterns using their variation in abundance (metagenomic datasets) and in activity (metatranscriptomic datasets) during the composting process. The correlation patterns derived from the abundance profile resulted in a graph composed by 40 nodes and 107 interactions, and four clusters (Fig. 5a). The correlation patterns derived from the activity profile resulted in a graph composed by 43 nodes and 76 interactions, and three clusters (Fig. 5b).

We observed a high concordance between both correlation analyses. Fifty-five interactions observed in the graph based on metagenomic datasets (Fig. 5a) are also present in the graph based on metatranscriptomic datasets (Fig. 5b). The correlations based on the activity of MAGs give us predictions on who are the key microbial players in the composting process as well as when and with whom they interact. In what follows we describe the main features of each activity cluster, highlighting high number of transcripts for selected cluster member MAGs on particular days (Supplementary Table S11).

Cluster 1: Seven MAGs form this cluster. Transcripts from members of this cluster are more abundant on initial days (D01 and D03) with a slight later increase on D64 (immediately after turning), followed by another increase on D99. Cluster members ZC4RG02 (g_Pseudoxanthomonas), ZC4RG04 (T.
bispora), and ZC4RG28 (f__Streptosporangiaceae) presented the highest number of transcripts in the initial days.

Cluster 2: This cluster is composed of 10 MAGs, all Firmicutes and mostly abundant and active between D3 and D15, followed by a peak on D64. Many transcripts from ZC4RG12 (g__Caldicoprobacter) and ZC4RG32 (s__Caldicoprobacter faecalis) related to lignocellulose degradation were identified, especially on D3, D7, and D15.

Cluster 3: This 26-MAG cluster is taxonomically diverse (it contains members of Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, and Proteobacteria phyla). The following cluster members are notable for expressing genes related to lignocellulose breakdown: ZC4RG13 (s__R. marinus), ZC4RG14 (f__Steroidobacteraceae), ZC4RG16 (s__T. fusca), ZC4RG26 (s__S. thermophilus), ZC4RG29 (f__Cyclobacteriaceae), ZC4RG36 (c__Anaerolinea), ZC4RG46 (o__Polyangiales), ZC4RG47 (g__Micromonospora), and ZC4RG48 (f__Roseiflexaceae); this activity is especially intense on D30, 78, and D99. ZC4RG20 (c__Gammaproteobacteria) and ZC4RG45 (s__T. agreste) express several genes associated with lignin degradation (i.e., annotated with CAZy AA families), especially on day 99.

Metabolic dependencies based on genome-scale models
Based on the results obtained with the co-occurrence analysis (Fig. 5b) and the activity of relevant functional genes (Supplementary Tables S7, S9, S10, S11, Fig. 2, and Supplementary Fig. S2), we identified MAGs according to their importance in the different stages of composting and the main functions associated with them (Fig. 6). We used this model in turn to assess the metabolic dependencies between these MAGs based on genome-scale models (Supplementary Fig. S3). The results revealed strong dependencies between some of the MAGs, as they received a maximum dependency score (Table 1). According to the models obtained, the most frequent compounds involved in the interactions between MAGs are hypoxanthine, H+\textsuperscript{+}, uracil, and phosphate. ZC4RG26 (s__Sphaerobacter thermophilus) had the highest level of dependency from other MAGs, predicted to be a metabolite receiver of 11 compounds, followed by ZC4RG08 (s__Pseudomonas thermotolerans) and ZC4RG20 (c__Gammaproteobacteria), predicted to receive six and five compounds, respectively.
ZC4RG04 (s__Thermobispora bispora) and ZC4RG22 (o__Luteitaleales) were predicted to have the highest number of interactions as metabolite producers. They were both predicted to be donors of seven compounds, followed by ZC4RG28 (f__Streptosporangiaceae), predicted as donor of six compounds. The set of keystone MAGs during the final composting stage (D99) presented higher possibilities for strong metabolite dependencies between MAGs compared to the other stages (Table 1).

In order to test if the predicted metabolic interactions are likely to be specific of the bacteria strains that were able to thrive in the composting microbiome, we chose three of the MAGs assigned to known species that were more frequently involved in the metabolic interactions model (ZC4RG26, ZC4RG08, ZC4RG04) and replaced them by the respective reference genomes from GenBank according to the taxonomic assignment based on GTDB: Sphaerobacter thermophilus (RefSeq: GCF_000024985.1), Pseudomonas thermotolerans (RefSeq GCF_000364625.1), and Thermobispora bispora (RefSeq GCF_000092645.1). The results show that several metabolic exchange interactions were lost (Fig. 7). For instance, in the models built using the RefSeq genomes, S. thermophilus and T. bispora were not able to share metabolites with each other. On the other hand, the RefSeq P. thermotolerans was able to share palmitate with other nodes, while ZC4RG08 (s__P. thermotolerans) was not (Fig. 7).

Discussion
Here we present detailed analyses of metagenome-assembled genomes (MAGs) from a composting process. A total of 60 environmental genomes of composting bacteria were obtained, of which 47 are potentially new bacterial species.

Among the recovered genomes, ZC4RG01 (s__Caldibacillus debilis), ZC4RG04 (s__T. bispora), ZC4RG13 (s__Rodothermus marinus), ZC4RG21 (s__Thermobifida fusca), ZC4RG26 (s__Sphaerobacter thermophilus), ZC4RG32 (s__Caldicoprobacter faecalis), and ZC4RG49 (s__[Clostridium] cellulosi) have been classified as species previously reported as being capable of cellulose degradation [11, 30–35]. Two Chloroflexi MAGs, ZC4RG36 (c__Anaerolineae) and ZC4RG48 (f__Roseiflexaceae), are additional lignocellulose degraders that we have found, having many CDSs
classified as CAZymes (326 and 313, respectively), exceeding the number of CAZymes in the much better-known lignocellulose degraders *T. bispora* [36] and *T. fusca* [37], corresponding to ZC4RG04 and ZC4RG21, with 174 and 150 CAZymes, respectively. Members from Chloroflexi have been reported in biomass-degrading environments using cultivation-independent methods, in some cases associated with the maturing phase of the composting process [35, 38, 39].

Lignocellulose breakdown depends on the efficiency of saccharification, which is often considered a bottleneck in the composting process [9]. Therefore, in addition to an analysis of MAGs and genes directly associated with biomass degradation, we also investigated other functions involved in microbial interactions and energy metabolism. We also carried out a co-occurrence analysis of MAGs based on their activity profiles during the process. Based on all these results, we propose here a framework for the molecular-microbial temporal dynamic in the composting process that we have studied (Fig. 6).

MAGs from clusters 1 and 2 (Fig. 5b) are the main constituents of the composting stage characterized by days 1, 3 (composting start), and 64 (recapitulation of composting start after the turning procedure) (Fig. 6). These MAGs represent bacterial populations expressing genes such as ARGs and classified with functions related to secondary metabolite production. These activities could be explained by intense competition between microorganisms. Indeed, these stages have high microbial diversity [3]. According to this interpretation, secondary metabolite production and ARG expression would be the consequence of the arms-shields race taking place in the composting microbial community [40–42]. Among the MAGs in these clusters, ZC4RG03 (*g__Calditerricola*) and ZC4RG07 (*g__Bacillus*) are the major producers of secondary metabolites. We hypothesize that these MAGs represent important bacterial players able to produce compounds with selective antimicrobial activity against pathogenic and opportunistic competitor bacteria; at the same time they seem able to consume easily degradable compounds, which we assume are particularly abundant during the initial composting stages (or right after the turning procedure). Niche protection through antagonistic competition has been observed in microbial systems associated with the human and animal gut [43–46].
According to our framework (Fig. 6), MAGs from cluster 2 are primarily active between days 1 and 15 (and on D64) (Supplementary Table S4). All of them are Firmicutes encoding hydrogenases. \( H_2 \)-oxidizing bacteria can use the molecular hydrogen produced during fermentative conversion of organic compounds [47], which possibly justifies the relevance of MAGs with hydrogenase activity in this stage. Hydrogenases participate in the mechanism that allows bacteria to store metabolic energy as an electrochemical potential across the membrane via the proton-motive force [48]. \( H_2 \) metabolism coupled with \( CO_2 \) as carbon source, which allows autotrophic growth via the acetyl coenzyme A (i.e., the Wood pathway) [49], is a metabolic strategy employed by acetogenic bacteria. These bacteria are strict anaerobes, most of which are also capable of heterotrophic growth [48].

In this context, Firmicutes MAGs ZC4RG12, 32, and 34 (Supplementary Table S1) (all belonging to cluster 2) would be the putative acetogens in the composting microbiome (Supplementary Table S10). Acetogens are ubiquitous in nature, including in thermophilic environments, and often constitute a fundamental group in the digestive tract of animals [48–52]. The latter could be the main source of acetogens for the composting microbiome analyzed here, given that a considerable portion of the composting material is made of feces from Zoo animals [3]. It has been reported that acetogens can outcompete methanogens and sulfate reducers given their higher metabolic flexibility [48, 53]. This could explain why these other two groups of bacteria were not detected among the recovered MAGs (Supplementary Table S10).

After day 15 we conjecture that easily degradable organic nutrients become more scarce [3], allowing other populations that can degrade more recalcitrant material to become dominant. Our results suggest that these populations are represented primarily by ZC4RG13 (s__R. marinus), ZC4RG22 (o__Luteitaleales), and ZC4RG29 (f__Cyclobacteriaceae), all from cluster 3 (Fig. 6). Around this time oxygen probably becomes more limited and denitrification processes come into play (Supplementary Fig. S2). Accordingly, the above MAGs, including ZC4RG26 (s__S. thermophilus), express CDSs annotated as nitrous-oxide reductase (including nosZ and nosD), which is the last step in denitrification, being therefore able to perform the complete pathway (Supplementary Table S10).
Anaerobic respiration using nitrous oxide ($N_2O$) is a widespread trait in prokaryotes, however not all denitrifiers encode this final step in denitrification [54]. Here we were able to detect four MAGs encoding nitrous-oxide reductase (Supplementary Table S10). One of them is an Acidobacteria (ZC4RG22) (also from cluster 3), and to our knowledge this is the first report of a nitrous-oxide reductase in this clade [54]. The ability to utilize nitrous oxide for respiration might be crucial to improve efficiency of nitrogen use in the composting process. In this context, it is worth mentioning that $N_2O$ is a potent greenhouse gas, and microbial conversion of $N_2O$ to $N_2$ is the unique sink known for $N_2O$ in the biosphere [55].

The following stages in the composting process (represented by days 30, 78 and 99) are mainly dominated by MAGs from cluster 3 that perform lignocellulose degradation, denitrification, and sulfur oxidation. In these stages there is a decrease in the overall phylogenetic microbial diversity [3]. Nevertheless, cluster 3 contains the largest and most diverse group of MAGs. We hypothesize that in these stages most nutrients derive from recalcitrant material (e.g. lignin). Examples of MAGs from cluster 3 are ZC4RG13 (s__R. marinus), ZC4RG08 (P. thermotolerans),, and ZC4RG20 (c__Gammaproteobacteria). ZC4RG20 has a high potential of AAs compared to other MAGs, and ZC4RG08 (s__P. thermotolerans) is the only one to have genes annotated as belonging to the AA1 family. AA1 enzymes that have been experimentally studied are multicopper oxidases that use diphenols and related substances as donors, with oxygen as acceptor, and known for their role in the enzymatic conversion of recalcitrant polysaccharides such as lignin [56].

By applying metabolic modeling methods, we were able to predict metabolic dependencies between MAGs (Supplementary Fig. S3). These results suggest that metabolic interactions in composting can be determined by complementary functions found in the genomes of producers and consumers (Supplementary Fig. S3 and Fig. 7). According to the Black Queen hypothesis, in order to increase fitness, one microbial population may lose genes related to a function when that function is already provided by another microbial population in the community. Therefore, the genomic differences between closely related strains are likely to be driven by local adaptation and coevolutionary
interactions [57, 58], which could explain why composting bacteria can present different metabolic dependencies compared with closely related strains that were isolated from other environments (Fig. 7).

Hypoxanthine and uracil are among the frequently exchanged compounds between keystone players (Table 1). It is known that many microbial groups, including members of the phylum Firmicutes, are auxotrophic for both purines and pyrimidines and rely on the salvage pathway for growth [59]. Dependencies on H⁺ exchange can be associated with hydrogenase activity (Supplementary Table S10) and proton flux across membranes, which is related to ATP synthesis, pH homeostasis, and maintenance of solute gradients [60]. A metabolite transported into the extracellular environment as a waste product by one bacterium is often used by neighboring bacteria [9, 58]. This could also explain the O² exchange flux predictions (Table 1). Oxygen can be a product derived from reactive oxygen species (ROS) detoxification systems, such as those encoded by chlorate-reducing bacteria via chlorite dismutase [61] (Supplementary Table S10). Due to intense redox activity, ROS-detoxification is a vital function in the composting microbiome, as observed also by the overall profile of dominant functions detected in the metatranscriptomic dataset (Supplementary Table S12).

Another important pathway detected in the activity profile is related to the phosphate starvation response (Supplementary Table S12), which is consistent with the frequency of metabolic dependencies based on phosphate exchange between MAGs (Table 1).

Conclusions
The results obtained here expand our knowledge of the taxonomic and functional diversity of composting bacteria. Based on our results we propose a model that predicts the key players in the composting process at each of the major stages. This model includes predicted metabolic interactions among the bacterial populations that we have identified, therefore providing unprecedented level of detail for microbial-molecular processes in composting.

Our results emphasize the importance of functions involved in efficient bioenergetic processes during biomass conversion, such as N₂O reduction, sulfur oxidation, and hydrogenases. Taken together, our results contribute to future research aiming at the engineering of efficient biomass-degrading
microbiomes.

Methods

**MAGs recovery from composting metagenomic data**

The composting metagenomic datasets on which this study is based have been described previously [3]. Briefly, the samples come from the composting facility at the State Zoological Park in the city of São Paulo, Brazil. Two composting cells were sampled: one called ZC3 and the other ZC4. For both, composting lasted 99 days. For ZC3, samples come from days 1, 30, 64, 78, and 99, and for ZC4 they come from days 1, 3, 7, 15, 30, 64, 67, 78, and 99). A turning procedure was performed on day 65 for ZC3 and on day 63 for ZC4. DNA shotgun sequencing was done for all samples, and metatranscriptomic sequencing was done for ZC4 samples.

Shotgun metagenomic reads from all samples were filtered and soft-trimmed (Q ≥ 12) using BBduk from the BBTools package (https://jgi.doe.gov/data-and-tools/bbtools/). Reads with length shorter than 80 bp were removed and the remaining reads were *de novo* assembled using Spades (k-mer = 21,33,55,77,99,113,121,127,—meta) [62]. To obtain MAGs, the following steps were carried out, for each composting cell (ZC3 and ZC4) (Supplementary Fig. S4): 1) reads from all samples were assembled and the contigs binned with Metabat2 [31]. 2) reads from individual samples were assembled and the corresponding contigs were binned using Metabat2. After these two steps, only bins with completeness at least 50% and contamination at most 10% were kept for further processing, based on results from CheckM [22]; 3) bins from individual samples (step 2) were compared with each other using Mash [63], which allowed us to establish when the “same” bin occurred on different days (Mash distance at most 0.05); 4) for each “distinct” bin determined in step 3 its reads were reassembled and the results rebinned with Metabat2 and MyCC [64]; 5) bins from step (1) and those from step (4) were compared, again using Mash; 6) the MAGs selected for additional analyses were those distinct bins with best completeness and contamination results (when there was more than one bin for the same MAG), provided completeness was at least 80% and contamination at most 11% (Supplementary Fig. S4 and Table S13).

**Taxonomic assignment**

MAG taxonomic assignment was based on GTDB [65, 66]. For those assignments that reached the
species level, we carried out further comparisons with reference genomes of those species (whenever available, with complete genomes). These comparisons were done with the ANI tool (http://enve-omics.ce.gatech.edu/ani/) and with GGDC [67]. In the text, we refer to MAGs by their identifiers, providing in parenthesis the GTDB classification, which uses the prefixes p, c, o, f, g, and s for phylum, class, order, family, genus, and species, respectively.

**Functional Annotation**

MAGs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline [68]. Their protein-coding gene sequences were compared against the Clusters of Orthologous Groups (COGs) [69] database using rpsblast+ [70], with a cut-off e-value of at most $10^{-5}$. COG categories were assigned to the best hits with cdd2cog script (https://github.com/aleimba/bac-genomics-scripts/tree/master/cdd2cog). The amino acid sequences of the predicted coding sequences (CDSs) were classified using the dbCAN HMM-based database for carbohydrate-active enzymes (CAZymes) [56]. CAZymes are categorized in different classes and families, including key enzymes for lignocellulose degradation such as glycoside hydrolases (GHs) and auxiliary activities (AAs), and the following complementary enzymes: glycosyltransferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs) and carbohydrate-binding modules (CBMs). The CDSs were also classified using a set of HMMs for detecting other metabolic pathways of interest [71], such as genes involved in denitrification, sulfur metabolism, hydrogen metabolism, and oxygen metabolism. Evidence supporting MAGs with strictly anaerobic metabolism was obtained based on the consistency between the results provided by TRAITAR and CDS classification as oxidases [72]. The global profile of functions in the metatranscriptomic dataset was determined using FMAP [73] (Supplementary Table S12).

We analyzed the presence of antibiotic resistance (AR) in our MAGs by comparing protein-coding sequences against the CARD database (April 2019) [74]. We filtered out all results below 70% identity and 85% sequence coverage.

We used antiSMASH (v.5.0) to find gene clusters involved in the biosynthesis of secondary metabolites [75].
Abundance and activity profiles of MAGs
The abundance profile of MAGs across the metagenomic datasets was obtained using the function quant_bins provided by metaWRAP followed by normalization based on TPM (transcripts per kilobase million) [76], which translates to genome copies per million reads in our context. Similarly, relative abundance of expressed CDSs was obtained by determining metatranscriptome reads that mapped to CDSs using BEDTools [77], followed by normalization based on TPM. We use the term transcripts to refer to MAG CDSs to which metatranscriptome reads could be mapped.

Co-occurrence of MAGs based on their abundance and activity profiles
In order to identify patterns of co-occurring bacteria represented by MAGs in the ZC4 datasets, we performed correlation analysis based on relative abundance of MAGs and their transcripts, as described. We used CONET [78] (Spearman $r^2 > 0.8$) and the resulting graphs were visualized in Cytoscape [79].

Metabolic interaction models
Based on MAG co-occurrence patterns and the relative abundance of transcripts with annotation related to biomass degradation, denitrification, sulfur metabolism, hydrogen metabolism, and oxygen metabolism, we defined subsets of MAGs according to their importance in the different stages of the composting process. For each subset we built a metabolic interaction model using SMETANA [80], based on genome-scale metabolic reconstructions that were obtained from files annotated in the PATRIC platform [81] (at the closest taxonomic level possible). The results were submitted to KBase [82] in order to run the Build Metabolic Model function, including the default option gapfilling, which relies on the ModelSEED Biochemistry Database [83]. With this method, metabolic genes were mapped onto biochemical reactions, and this information was integrated with information on stoichiometry reactions, subcellular localization, biomass composition, and estimation of thermodynamic feasibility, in order to produce a detailed stoichiometric model of metabolism at the genome scale. Metabolic dependency score calculated by SMETANA is normalized to a range between 0 and 1, meaning complete independency and complete dependency, respectively [80].

Declarations
Ethics approval and consent to participate
Consent for publication
Not applicable

Availability of data and material
The sequence of all MAGs described in this work are available from GenBank, and the accession numbers are listed in Supplementary Table S1.

Competing interests
The authors declare that they have no competing interests.

Funding
Funding for this research was provided in part by grant 2011/50870–6 from the São Paulo Research Foundation (FAPESP). Additional funding was provide by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (the BIGA project). LMSM and RVP were supported by fellowships from CAPES. LPPB was supported by FAPESP fellowship 18/19247–0. AMDS and JCS are supported in part by Research Fellowship Awards from the National Council for Scientific and Technological Development (CNPq). The funders had no role in study design, data collection, analysis, decision to publish or preparation of the manuscript.

Authors’ contributions
JCS and AMDS conceived the study. LPPB, RVP, LFM, and LMSM performed most analyses. JSLP did the phylogenetic analysis and FBS did the comparison between MAGs in this study and other genomes. All authors participated in discussions of the results. LPPB, RVP, LFM, and LMSM contributed to the writing, and AMDS and JCS wrote the final manuscript.

Acknowledgements
The authors are grateful to Paulo Bressan, João Batista da Cruz, and Patricia Locosque Ramos for support to this project and for allowing sampling at the São Paulo Zoo Compost Facility; and to Carlos Morais Piroupo for computational support.

References
1. Ryckeboer J, Mergaert J, Vaes K, Klammer S, de Clercq D, Coosemans J, Insam H, Swings J (2003) A survey of bacteria and fungi occurring during composting and self-heating processes. Ann Microbiol 53:349–410
2. Martins LF, Antunes LP, Pascon RC, et al (2013) Metagenomic analysis of a tropical composting operation at the são paulo zoo park reveals diversity of biomass degradation functions and organisms. PLoS One 8:e61928

3. Antunes LP, Martins LF, Pereira RV, et al (2016) Microbial community structure and dynamics in thermophilic composting viewed through metagenomics and metatranscriptomics. Sci Rep 6:38915

4. Kolinko S, Wu Y-W, Tachea F, et al (2018) A bacterial pioneer produces cellulase complexes that persist through community succession. Nat Microbiol 3:99-107

5. Jurado M, López MJ, Suárez-Estrella F, Vargas-García MC, López-González JA, Moreno J (2014) Exploiting composting biodiversity: study of the persistent and biotechnologically relevant microorganisms from lignocellulose-based composting. Bioresour Technol 162:283-293

6. Cerda A, Artola A, Font X, Barrena R, Gea T, Sánchez A (2018) Composting of food wastes: Status and challenges. Bioresour Technol 248:57-67

7. Viikari L, Vehmaanperä J, Koivula A (2012) Lignocellulosic ethanol: From science to industry. Biomass Bioenergy 46:13-24

8. Lynd LR (2017) The grand challenge of cellulosic biofuels. Nat Biotechnol 35:912-915

9. Jiménez DJ, Dini-Andreote F, DeAngelis KM, Singer SW, Salles JF, van Elsas JD (2017) Ecological Insights into the Dynamics of Plant Biomass-Degrading Microbial Consortia. Trends Microbiol 25:788-796

10. Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016) Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. Biotechnol Biofuels 9:22

11. Lemos LN, Pereira RV, Quaggio RB, Martins LF, Moura LMS, da Silva AR, Antunes LP, da Silva AM, Setubal JC (2017) Genome-Centric Analysis of a Thermophilic and
12. Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N (2013) Changes in bacterial and fungal communities across compost recipes, preparation methods, and composting times. PLoS One 8:e79512

13. Zainudin MHM, Hassan MA, Shah UKM, Abdullah N, Tokura M, Yasueda H, Shirai Y, Sakai K, Baharuddin AS (2014) Bacterial community structure and biochemical changes associated with composting of lignocellulosic oil palm empty fruit bunch. BioResources 9:316–335

14. Danon M, Franke-Whittle IH, Insam H, Chen Y, Hadar Y (2008) Molecular analysis of bacterial community succession during prolonged compost curing. FEMS Microbiol Ecol 65:133–144

15. Sorek N, Yeats TH, Szemenyei H, Youngs H, Somerville CR (2014) The Implications of Lignocellulosic Biomass Chemical Composition for the Production of Advanced Biofuels. Bioscience 64:192–201

16. Montella S, Ventorino V, Lombard V, Henrissat B, Pepe O, Faraco V (2017) Discovery of genes coding for carbohydrate-active enzyme by metagenomic analysis of lignocellulosic biomasses. Sci Rep 7:42623

17. Lindemann SR, Bernstein HC, Song H-S, Fredrickson JK, Fields MW, Shou W, Johnson DR, Beliaev AS (2016) Engineering microbial consortia for controllable outputs. ISME J 10:2077–2084

18. D'haeseleer P, Gladden JM, Allgaier M, et al (2013) Proteogenomic analysis of a thermophilic bacterial consortium adapted to deconstruct switchgrass. PLoS One 8:e68465

19. Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P, Hirsch PR, Vogel TM (2011) Accessing the soil metagenome for studies of microbial diversity. Appl
Environ Microbiol 77:1315–1324

20. Castañeda LE, Barbosa O (2017) Metagenomic analysis exploring taxonomic and functional diversity of soil microbial communities in Chilean vineyards and surrounding native forests. PeerJ 5:e3098

21. Chopyk J, Allard S, Nasko DJ, Bui A, Mongodin EF, Sapkota AR (2018) Agricultural Freshwater Pond Supports Diverse and Dynamic Bacterial and Viral Populations. Front Microbiol 9:792

22. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055

23. Sangwan N, Xia F, Gilbert JA (2016) Recovering complete and draft population genomes from metagenome datasets. Microbiome 4:8

24. Parks DH, Rinke C, Chuvchchina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol 2:1533–1542

25. Stewart RD, Auffret MD, Warr A, et al (2018) Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. Nat Commun 9:870

26. Campanaro S, Treu L, Kougias PG, Luo G, Angelidaki I (2018) Metagenomic binning reveals the functional roles of core abundant microorganisms in twelve full-scale biogas plants. Water Res 140:123–134

27. Garza DR, Dutilh BE (2015) From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. Cell Mol Life Sci 72:4287-4308

28. Anantharaman K, Brown CT, Hug LA, et al (2016) Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 7:13219
29. Bowers RM, Kyrpides NC, Stepanauskas R, et al (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731

30. Gomez del Pulgar EM, Saadeddin A (2014) The cellulolytic system of Thermobifida fusca. Crit Rev Microbiol 40:236–247

31. Bjornsdottir SH, Blondal T, Hreggvidsson GO, Eggertsson G, Petursdottir S, Hjorleifsdottir S, Thorbjarnardottir SH, Kristjansson JK (2006) Rhodothermus marinus: physiology and molecular biology. Extremophiles 10:1–16

32. Hiras J, Wu Y-W, Deng K, et al (2016) Comparative Community Proteomics Demonstrates the Unexpected Importance of Actinobacterial Glycoside Hydrolase Family 12 Protein for Crystalline Cellulose Hydrolysis. MBio. https://doi.org/10.1128/mBio.01106–16

33. Yokoyama H, Wagner ID, Wiegel J (2010) Caldicoprobacter oshimai gen. nov., sp. nov., an anaerobic, xylanolytic, extremely thermophilic bacterium isolated from sheep faeces, and proposal of Caldicoprobacteraceae fam. nov. Int J Syst Evol Microbiol 60:67–71

34. Zhang K, Chen X, Schwarz WH, Li F (2014) Synergism of glycoside hydrolase secretomes from two thermophilic bacteria cocultivated on lignocellulose. Appl Environ Microbiol 80:2592–2601

35. Wang Z, Cao G, Zheng J, Fu D, Song J, Zhang J, Zhao L, Yang Q (2015) Developing a mesophilic co-culture for direct conversion of cellulose to butanol in consolidated bioprocess. Biotechnol Biofuels 8:84

36. Liolios K, Sikorski J, Jando M, et al (2010) Complete genome sequence of Thermobispora bispora type strain (R51). Stand Genomic Sci 2:318–326

37. Gomez del Pulgar EM, Saadeddin A (2014) The cellulolytic system of Thermobifida...
38. Wong MT, Wang W, Lacourt M, Couturier M, Edwards EA, Master ER (2016) Substrate-Driven Convergence of the Microbial Community in Lignocellulose-Amended Enrichments of Gut Microflora from the Canadian Beaver (Castor canadensis) and North American Moose (Alces americanus). Front Microbiol 7:961

39. Puentes-Téllez PE, Falcao Salles J (2018) Construction of Effective Minimal Active Microbial Consortia for Lignocellulose Degradation. Microb Ecol 76:419–429

40. D’Costa VM, Griffiths E, Wright GD (2007) Expanding the soil antibiotic resistome: exploring environmental diversity. Curr Opin Microbiol 10:481–489

41. Nesme J, Simonet P (2015) The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. Environ Microbiol 17:913–930

42. Blanco P, Hernando-Amado S, Reales-Calderon JA, Corona F, Lira F, Alcalde-Rico M, Bernardini A, Sanchez MB, Martinez JL (2016) Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. Microorganisms. https://doi.org/10.3390/microorganisms4010014

43. Hsiao A, Ahmed AMS, Subramanian S, Griffin NW, Drewry LL, Petri WA Jr, Haque R, Ahmed T, Gordon JL (2014) Members of the human gut microbiota involved in recovery from Vibrio cholerae infection. Nature 515:423–426

44. Iacob S, Iacob DG, Luminos LM (2018) Intestinal Microbiota as a Host Defense Mechanism to Infectious Threats. Front Microbiol 9:3328

45. Mousa WK, Athar B, Merwin NJ, Magarvey NA (2017) Antibiotics and specialized metabolites from the human microbiota. Nat Prod Rep 34:1302–1331

46. Garcia-Gutierrez E, Mayer MJ, Cotter PD, Narbad A (2019) Gut microbiota as a source of novel antimicrobials. Gut Microbes 10:1–21
47. Hungate RE (1967) Hydrogen as an intermediate in the rumen fermentation. Arch Mikrobiol 59:158-164

48. Schuchmann K, Müller V (2016) Energetics and Application of Heterotrophy in Acetogenic Bacteria. Appl Environ Microbiol 82:4056-4069

49. Drake HL, Gössner AS, Daniel SL (2008) Old acetogens, new light. Ann N Y Acad Sci 1125:100-128

50. Kamlage B, Gruhl B, Blaut M (1997) Isolation and characterization of two new homoacetogenic hydrogen-utilizing bacteria from the human intestinal tract that are closely related to Clostridium coccoides. Appl Environ Microbiol 63:1732-1738

51. Drake HL, Daniel SL, Küsel K, Matthies C, Kuhner C, Braus-Stromeyer S (1997) Acetogenic bacteria: what are the in situ consequences of their diverse metabolic versatilities? Biofactors 6:13-24

52. Gagen EJ, Denman SE, Padmanabha J, Zaduke S, Al Jassim R, Morrison M, McSweeney CS (2010) Functional gene analysis suggests different acetogen populations in the bovine rumen and tammar wallaby forestomach. Appl Environ Microbiol 76:7785-7795

53. Lever MA (2011) Acetogenesis in the energy-starved deep biosphere - a paradox? Front Microbiol 2:284

54. Hallin S, Philippot L, Löffler FE, Sanford RA, Jones CM (2017) Genomics and Ecology of Novel N2O-Reducing Microorganisms. Trends Microbiol. https://doi.org/10.1016/j.tim.2017.07.003

55. Hallin S, Philippot L, Löffler FE, Sanford RA, Jones CM (2018) Genomics and Ecology of Novel N2O-Reducing Microorganisms. Trends Microbiol 26:43-55

56. Lombard V, Golaonda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490-5
57. Thompson JN (2013) Relentless Evolution. In: University of Chicago Press. https://www.press.uchicago.edu/ucp/books/book/chicago/R/bo15112984.html. Accessed 13 Feb 2020

58. Jeffrey Morris J (2015) Black Queen evolution: the role of leakiness in structuring microbial communities. Trends Genet 31:475–482

59. Kilstrup M, Hammer K, Ruhdal Jensen P, Martinussen J (2005) Nucleotide metabolism and its control in lactic acid bacteria. FEMS Microbiol Rev 29:555–590

60. Kakinuma Y (1998) Inorganic cation transport and energy transduction in Enterococcus hirae and other streptococci. Microbiol Mol Biol Rev 62:1021–1045

61. Ettwig KF, Speth DR, Reimann J, Wu ML, Jetten MSM, Keltjens JT (2012) Bacterial oxygen production in the dark. Front Microbiol 3:273

62. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834

63. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM (2016) Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17:132

64. Lin H-H, Liao Y-C (2016) Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. Sci Rep 6:24175

65. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarszewski A, Chaumeil P-A, Hugenholtz P (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004

66. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. https://doi.org/10.1093/bioinformatics/btz848
67. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60

68. Tatusova T, DiCuccio M, Badretdin A, Chetverinin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J (2016) NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614-6624

69. Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43:D261-9

70. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421

71. Matheus Carnevali PB, Schulz F, Castelle CJ, et al (2019) Hydrogen-based metabolism as an ancestral trait in lineages sibling to the Cyanobacteria. Nat Commun 10:463

72. Weimann A, Mooren K, Frank J, Pope PB, Bremges A, McHardy AC (2016) From Genomes to Phenotypes: Traitar, the Microbial Trait Analyzer. mSystems. https://doi.org/10.1128/mSystems.00101–16

73. Kim J, Kim MS, Koh AY, Xie Y, Zhan X (2016) FMAP: Functional Mapping and Analysis Pipeline for metagenomics and metatranscriptomics studies. BMC Bioinformatics 17:420

74. Jia B, Raphenya AR, Alcock B, et al (2017) CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 45:D566–D573

75. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal
76. Uritskiy GV, DiRuggiero J, Taylor J (2018) MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 6:158

77. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842

78. Faust K, Raes J (2016) CoNet app: inference of biological association networks using Cytoscape. F1000Res 5:1519

79. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504

80. Zeleznjak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR (2015) Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A 112:6449–6454

81. Wattam AR, Davis JJ, Assaf R, et al (2017) Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535–D542

82. Arkin AP, Cottingham RW, Henry CS, et al (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569

83. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28:977–982

84. Banat IM, Marchant R, Rahman TJ (2004) Geobacillus debilis sp. nov., a novel obligately thermophilic bacterium isolated from a cool soil environment, and reassignment of Bacillus pallidus to Geobacillus pallidus comb. nov. Int J Syst Evol Microbiol 54:2197–2201
85. Berendsen EM, Wells-Bennik MHJ, Krawczyk AO, de Jong A, van Heel A, Holsappel S, Eijlander RT, Kuipers OP (2016) Draft Genome Sequences of Seven Thermophilic Spore-Forming Bacteria Isolated from Foods That Produce Highly Heat-Resistant Spores, Comprising Geobacillus spp., Caldibacillus debilis, and Anoxybacillus flavithermus. Genome Announc. https://doi.org/10.1128/genomeA.00105–16

86. Henssen A (1957) Beiträge zur Morphologie und Systematik der thermophilen Actinomyceten. Archiv für Mikrobiologie 26:373-414

87. Liolios K, Sikorski J, Jando M, et al (2010) Complete genome sequence of Thermobispora bispora type strain (R51). Stand Genomic Sci 2:318-326

88. Steger K, Jarvis A, Vasara T, Romantschuk M, Sundh I (2007) Effects of differing temperature management on development of Actinobacteria populations during composting. Res Microbiol 158:617-624

89. Bellamy WD (1977) Cellulose and lignocellulose digestion by thermophilic actinomyces for single-cell protein production. Developments in Industrial Microbiology 18:

90. Hatayama K, Shoun H, Ueda Y, Nakamura A (2005) Planifilum fimeticola gen. nov., sp. nov. and Planifilum fulgidum sp. nov., novel members of the family “Thermoactinomycetaceae” isolated from compost. Int J Syst Evol Microbiol 55:2101–2104

91. Manaia CM, Moore ERB (2002) Pseudomonas thermotolerans sp. nov., a thermotolerant species of the genus Pseudomonas sensu stricto. Int J Syst Evol Microbiol 52:2203-2209

92. Alfredsson GA, Kristjansson JK, Hjörleifsdottir S, Stetter KO (1988) Rhodothermus marinus, gen. nov., sp. nov., a Thermophilic, Halophilic Bacterium from Submarine Hot Springs in Iceland. Microbiology 134:299-306
93. Demharter W, Hensel R, Smida J, Stackebrandt E (1989) Sphaerobacter thermophilus gen. nov., sp. nov. A Deeply Rooting Member of the Actinomycetes Subdivision Isolated from Thermophilically Treated Sewage Sludge. Syst Appl Microbiol 11:261-266

94. Schroder K-H, Naumann L, Kroppenstedt RM, Reischl U (1997) Mycobacterium h assiacum sp. nov., a New Rapidly Growing Thermophilic Mycobacterium. Int J Syst Bacteriol 47:86-91

95. Jiang SH, Roberts DM, Clayton PA, Jardine M (2013) Non-tuberculous mycobacterial PD peritonitis in Australia. Int Urol Nephrol 45:1423-1428

96. Deinhardt-Emmer S, Höring S, Mura C, Hillemann D, Hermann B, Sachse S, Bohnert J, Löffler B (2018) First Time Isolation of Mycobacterium h assiacum From a Respiratory Sample. Clin Med Insights Circ Respir Pulm Med 12:1179548417747529

97. Korn-Wendisch F, Rainey F, Kroppenstedt RM, Kempf A, Majazza A, Kutzner HJ, Stackebrandt E (1995) Thermocrispum gen. nov., a new genus of the order Actinomycetales, and description of Thermocrispum municipale sp. nov. and Thermocrispum agreste sp. nov. Int J Syst Bacteriol 45:67-77

98. He YL, Ding YF, Long YQ (1991) Two cellulolytic Clostridium species: Clostridium cellulosi sp. nov. and Clostridium cellulofermentans sp. nov. Int J Syst Bacteriol 41:306-309

99. Hou L, Dutta SK (2000) Phylogenetic characterization of several para- and meta-PCB dechlorinating Clostridium species: 16s rDNA sequence analyses. Lett Appl Microbiol 30:238-243

Table
Due to technical limitations, Table 1 is provided in the Supplementary Files section.

Figures
Figure 1

Metabolic potential of MAGs from the ZC4 composting cell based on CAZymes. (a) CAZyme genes in 14 MAGs with at least 40 genes annotated as GHs (dashed line). (b) Breakdown of GH families for the same 14 MAGs as in (a). (c) Comparison of numbers of genes annotated as GHs related to lignocellulose degradation in the top six degraders (ZC4RG13, ZC4RG29, ZC4RG32, ZC4RG36, ZC4RG46, and ZC4RG48).
Figure 2

Heatmap representing the abundance of transcripts associated with CDSs annotated with functions related to lignocellulose degradation. The scale in shades of green is based on relative abundance (TPM).
Secondary metabolite cluster types detected among the ZC4 MAGs. The X-axis represents the number of clusters detected. The colors in the bars correspond to phylum assignment of MAGs.
Figure 4

Antibiotic resistance genes (ARGs) profile in ZC4 MAGs.
MAGs co-occurrence based on their relative abundances in (a) metagenomes and (b) metatranscriptomes. Nodes represent MAGs and edges represent correlations ($r^2 > 0.8$) based on the abundance profile in the datasets across composting. Different shapes indicate different Phyla. Different colors indicate different Class.
Schematic representation of keystone microbial players according to their importance in the different stages of composting. MAGs are represented as roughly circular numbered shapes, and their colors reflect the cluster they were assigned to (Fig 5b and Supplementary Table S4). Lignocellulose breakdown and relevant active functions are represented across the stages by the various symbols, and the irregular background shapes connect MAGs that express the same functions. The turning procedure was performed on day 63, therefore day 64 is considered the recapitulation of composting start, based on the patterns of microbial functions and activity that we observed in the present study and in our previous work [3].
Figure 7

Major metabolic dependencies between genomes of known bacteria species in the composting process studied. ZC4RG04, ZC4RG08, and ZC4RG26 were identified as the known bacteria species presenting major levels of metabolic dependencies (Table 1). The figure represents all the strong metabolic interactions (i.e., maximum dependency score) predicted for these MAGs (a). In panel (b) these MAGs were replaced by their closest relatives in the NCBI RefSeq database, and the consequent predicted changes in metabolic interactions are denoted with fading colors and by dotted arrows.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

Supplementary_Tables2020-03-22.xlsx
Supplementary_Figures2020-03-22.docx
Table_1.xlsx