Wheat domestication has dramatically affected the root-associated microbiome

Marta Kinnunen-Grubb  
Novozymes A/S

Rumakanta Sapkota  
Aarhus Universitet

Marta Vignola  
University of Glasgow

Ines Marques Nunes  
Novozymes A/S

Mogens Nicolaisen (mn@agro.au.dk)  
Aarhus Universitet  https://orcid.org/0000-0002-0407-2488

Research

Keywords: Domestication, Microbiome, Root, Wheat, Bacteria, Fungi, Selective pressure

DOI: https://doi.org/10.21203/rs.3.rs-20588/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract
Background: Plants-microbiome associations are the results of millions of years of co-evolution. Due to the accelerated plant evolution during domestication of crops and aided by cultivation in non-native highly managed soils, plant-microbe links created through co-evolution could have been lost. Therefore, we hypothesize that dramatic effects on the root-associated microbiome occurred during domestication of wheat.

Results: To uncover domestication effects we analyzed root associated fungal and bacterial communities in a transect of wheat evolution including modern wheat cultivars, landraces, Triticum aestivum ssp. spelta and ancestors of wheat including T. turgidum ssp. dicoccum, T. monococcum ssp. monococcum and T. monococcum ssp. aegilopoides at three growth phases shortly after emergence of seedlings. We found that numbers of observed microbial taxa were highest in the landraces, which had been domesticated in low-input agricultural systems, and lowest in wheat ancestors that evolved in native soils. The root-associated microbial community of modern cultivars was significantly different from that of landraces and ancestors of wheat. Old wheat accessions enriched Acidobacteria and Actinobacteria, while modern cultivars enriched OTUs from Candidatus Saccharibacteria, Verrucomicrobia and Firmicutes. The fungal pathogens Fusarium, Neoascochyta and Microdochium were enriched in modern cultivars. The composition of root-associated microbial communities of modern wheat cultivars significantly followed patterns predicted by the neutral community assembly model. Our observations allowed us to suggest that a stronger selective pressure drives the root-associated microbiome of ancient wheat accessions than that of modern wheat cultivars.

Conclusions: Here we demonstrate that the wheat root-associated microbiome has dramatically changed through a transect of evolution from wheat ancestors over landraces to modern cultivars. Colonization of roots of ancient accessions was slower than in modern cultivars, and the root-associated microbiome of ancient wheat accessions was driven by stronger selective pressure than that of the modern wheat cultivars. We identified several taxa including Acidobacteria and Actinobacteria enriched in old cultivars and fungal wheat pathogens that were enriched in modern cultivars.

Background
Plants are associated with massive numbers of microorganisms from a vast array of taxa, and the rhizosphere is considered one of the most active microbiological environments on Earth [1]. The associations between the plant root and its microbiome are the result of millions of years of co-evolution [2], and it is well documented that host genotype has significant, albeit minor, effects on microbial community composition, both aboveground [3] and belowground [4, 5]. Many plant-associated microorganisms are crucial for plant growth, nutrient acquisition, as well as protection against biotic and abiotic stresses [6].
In contrast to natural evolution, domestication is a fast anthropological selection of crops, which, like natural selection, is based on genetic diversity and selection of desired traits. Modern breeding has further accelerated the domestication process with a focus on selection of fertilizer-responsive high-yielding and disease-resistant cultivars. Thus, crop domestication has occurred in two distinct stages: an early stage characterized by a selection of the best performing varieties, and a later stage characterized by a targeted breeding for specific plant traits. In parallel, agriculture has undergone dramatic intensification during the last century with increasing agronomic management including higher levels of fertilizer and pesticide inputs. Considering domestication is a result of human-induced selection rather than natural evolution, our understanding of domestication-induced evolutionary interactions between microorganisms and plant traits is limited.

It is evident that domestication caused changes of the plant genome through hybridization events and enrichment of genes in the process of selection of desirable plant traits. Such selection pressure for both yield and aboveground traits, in combination with shifts towards high-input agriculture and neglect of root processes and performance, may inadvertently have led to depletion of members of the root microbiota that were important for nutrient acquisition and pathogen control in natural low-input conditions [7]. In common sunflower (Helianthus annuus) Leff et al. [8] demonstrated how domestication affected fungal communities, but not bacterial communities, and they found that modern sunflower varieties had a lower relative abundance of putative fungal pathogens. Similarly, bacterial isolates from sugar beet roots (Beta vulgaris ssp. vulgaris) were more active against phytopathogens, compared to isolates from the wild ancestor (B. vulgaris ssp. In contrast, the rhizomicrobiome of wild rice was less sensitive to introduction of a fungal pathogen and an introduced pathogen had lower abundance in wild rice [10]. In barley, a small but significant effect of domestication was manifested mainly on the abundance of several OTUs from various taxa, rather than on single OTUs [11]. In contrast, domestication effects were abundant in foxtail millet (Setaria italica) and its wild ancestor (S. viridis), including effects on Betaproteobacteria and Firmicutes, which were enriched in S. italica [12]. In wild and domesticated Phaseolus vulgaris, a gradual decrease in relative abundance of members of Bacteroidetes, and an increase of members of Actinobacteria and Proteobacteria was observed from wild to modern accessions [13]. Bouffaud et al. [14] found that the evolutionary history of Poaceae (maize, sorghum, wheat and teosinte) was shaping bacterial communities, and that increases in phylogenetic distance were reflected in increasing differences between microbial communities.

Modern wheat cultivars are the product of at least 10,000 years of human selection during breeding and cultivation. Modern bread wheat (Triticum aestivum L.) has an allo-hexaploid genome consisting of three subgenomes (AABBDD) resulting from hybridization events between T. urartu (AA genome) and a close relative to Aegilops speltoides (BB genome), and a later hybridization with the wild diploid A. tauschii (DD genome) [15]. T. turgidum ssp. dicoccoides and T. turgidum ssp. dicoccum are allo-tetraploid emmer wheat varieties (BBAA genomes) suggested as direct progenitors of modern wheat [16]. Domesticated emmer, T. turgidum ssp. dicoccum, was first cultivated in the fertile crescent around 10,000 BCE. Einkorn, T. monococcum ssp. monococcum and T. monococcum ssp. aegilopoides are diploid species with AA genomes, also cultivated in the fertile crescent [17]. Spelt wheat (T. triticum ssp. spelta) represents a
hexaploid (AABBDD) type of hulled wheat, characterized by adaptation to a wide range of environments [18]. Several wheat traits have changed dramatically during domestication such as root architecture and exudation of primary and secondary metabolites [19–21]. This may have had profound effects on root and rhizosphere microbial communities, as demonstrated in a study of plant growth-promoting bacteria associated with roots of ancient and modern wheat [22]. It has further been shown that modern wheat types are less dependent on arbuscular mycorrhizae than older types and that old accessions benefit more from mycorrhizal symbiosis [23]. It was suggested that this could be caused by the highly fertile conditions of soils during breeding and cultivation of modern cultivars [24].

It is well known that the soil surrounding the roots is the major source of root-associated microbes. Understanding assembly processes of root associated microbial communities, and their diversity and ecology may contribute to future breeding programs. However, our knowledge of these processes in a plant evolutionary context is limited. In the present study, we hypothesized that root-associated microbial community composition changed with wheat domestication. To test this hypothesis, we selected three distinct genetic groups of Triticum depicting two major events in domestication history. The genetically oldest group that we chose was Triticum accessions of T. monococcum (AA genome; Tm) and T. turgidum (BBAA genome; Tt). To represent the wheat T. aestivum (genome BBAADD), we included landraces grown before the 1940s and one accession of T. triticum ssp. spelta. Four commercial cultivars currently grown in Denmark were used as representing modern wheat. We grew those accessions under field conditions in agricultural soils and compared root fungal and bacterial communities in field plots at three time points shortly after emergence of the seedlings in the autumn.

**Results**

**Amplicon sequencing quality**

Sequencing of the bacterial 16S rRNA amplicon library resulted in 8,806 OTUs in 210 samples (excluding controls). Number of reads per sample ranged between 20,000 and 40,000 (Figure S1). Samples with less than 2,000 reads were removed from downstream analysis (n=16). Furthermore, 504 OTUs unclassified at kingdom level and 29 sequences classified as chloroplast at family level were removed from the dataset. After clustering, the length of representative sequences of 16S V3-V4 region was 417 ± 10 bp (mean ± sd). Therefore, 11 OTUs with sequences shorter than 390 bp were eliminated from the dataset (Figure S1). Sequencing of the fungal ITS amplicon library resulted in 561 OTUs in 210 samples (excluding controls). Number of reads per samples ranged between 30,000 and 40,000 (Figure S1). Samples with less than 2000 reads were removed from the downstream analysis (n=10). Furthermore, 170 sequences not classified at kingdom level were removed from the analysis. Fourteen sequences smaller than 135 bp were removed from the dataset (Figure S1).

Rarefaction curves indicated that the number of reads sampled the variation in fungal communities whereas bacterial variation was not covered completely in our sampling strategy (Figure S2).

**Community alpha diversity**
As expected, bulk soil alpha diversity was always higher than the alpha diversity of the root samples (Figure 1 and Table S3 and S4). Alpha diversity in roots increased from BBCH10 to BBCH21. Differences in alpha diversity among genetic groups, measured as observed species richness and Shannon diversity indices, were most pronounced at BBCH 10, whereas the alpha diversity approached similar levels at BBCH21. For both bacteria and fungi, Landraces + Tas had the highest alpha diversity of all accessions. Tt generally had the lowest alpha diversity, most clearly observed in the bacterial dataset.

**Microbial communities separate according to wheat evolution**

Comparative analysis of modern cultivars, Landraces + Tas, and ancient accessions (Tm and Tt) separated those wheat groups in NMDS plots, but differently at the three growth stages. This indicates a significant effect of wheat genotypes and BBCH on the microbial community composition (p=0.001 for each BBCH) (Figure 2). We observed clear separation of bulk soil microbial communities from the root communities of all wheat accessions, the latter being more dynamic over time. Although the root bacterial and fungal community composition shifted between growth stages, the separation of modern cultivars from Landraces + Tas and Tt and Tm was remarkable. We observed noticeable differences between the wheat genetic groups. Landraces + Tas clearly clustered at all growth stages while the microbial community composition of the roots of Tt became more similar over time and approached the other genetic groups, except the modern cultivars, indicating a stronger succession of microbial communities in the Tt roots than in the roots of T. aestivum.

Comparing the dynamics of bacterial and fungal communities, we observed a stronger selection on bacterial communities compared to fungal communities at the first growth stage (based on stronger clustering of samples). However, over time, selection in the fungal communities increased. Fungal community composition in roots of ancient accessions was distinct from modern cultivars, and bacterial communities of modern cultivars were even more clearly separated from the landrace/ancient lines. Furthermore, overall fungal community composition in roots was similar to the bulk soil community composition, whereas the bacterial community composition in bulk soils clearly separated from root communities. PERMANOVA tests confirmed that both fungal and bacterial communities in soil and roots were significantly different in the total dataset (p<0.001), confirming our observations of the NMDS plots (Table 2). When root associated microbial data was split according to growth stage, the different wheat groups had significantly different communities (p<0.001; fungi at BBCH21, p<0.01). However, the wheat genetic groups explained a higher proportion of the variation in bacterial communities.

**Taxa significantly enriched in the roots of old and modern wheat lines**

For this analysis, we grouped the wheat lines into old (Tm, Tt, landraces and Tas) and modern lines (KWS Desanto, Substance, Sherif and Torp). We observed differences in taxa enriched in the roots of old and modern communities (Figure 3, Figure S3-S5). The most remarkable differences in bacterial enrichment could be accounted to OTUs belonging to Acidobacteria and Actinobacteria primarily in old cultivars at BBCH21, while modern cultivars enriched OTUs from Saccharibacteria, Verrucomicrobia and Firmicutes (Figure 3a). Furthermore, modern cultivars had higher numbers of enriched OTUs of Bacteroidetes at...
BBCH10, but this effect was absent at BBCH21. We also observed phylum level differences of fungal taxa enriched in the roots (Figure 3b). Taxa belonging to Ascomycota and Basidiomycota were enriched in both old and modern cultivars, but the number of enriched Basidiomycota was much higher in the old wheat lines. Although at a low number (and abundance), Mortierellomycota, Olpidiomycota and Chytridiomycota taxa were solely enriched in the roots of old cultivars. At OTU level, we observed a stronger enrichment of bacterial families in the modern cultivars than in ancient lines (Figure S5a). Interestingly, similar patterns were not observed for the fungal communities (Figure S5b). Modern wheat cultivars strongly enriched for OTUs belonging to Rhizobiaceae, Flavobacteriaceae and Pedobacter. Sphingomonadaceae was the most significantly enriched family in the communities of the old lines (Figure S5a). Microdochium, Neoascocytta, Sporobolomyces and Fusarium were enriched in modern cultivars whereas basidiomyceteous yeasts such as Filobasidium, Holtermaniella and Psathyrella, and the ascomycete Aureobasidium were enriched in the old lines (Figure S5b). Surprisingly, we observed that genera within Basidiomycota were more frequently enriched (and with higher significance) in old lines, while Ascomycota dominated in the modern cultivars.

**What drives the community assembly in the roots of wheat?**

Bacterial communities from both modern and old wheat accessions showed very strong correlation with the neutral model (NCM) [25] when the bulk soil community was considered as the source (Figure 4). The model explained 84% and 87% (Pearson correlation) of the observed variation in the frequencies of OTUs present in the local bacterial communities associated with modern and old lines, respectively. Communities associated with the modern cultivars were found to have a higher number of OTUs following the neutral prediction than the old lines. Indeed, 90% of the OTUs associated with modern cultivars were shared with the source (Table 3) and, among these, 75% followed NCM. Communities from old wheat accessions shared a lower percentage of OTUs with the source (86%), with only 61% being neutrally assembled (Table 3). Fungal communities showed a similar pattern. Communities from modern and old lines both correlated strongly with the model (0.84 Pearson correlation for both). They both shared a high percentage of OTUs with the source (93% for the modern and 91% for the old) with 69% and 66% of these OTUs following the neutral model respectively.

NRI (net relatedness index) values were calculated for the communities associated with the three distinct environments. Average values were 7.03±2.5, 2.16±0.78 and 1.4±1.08 for bulk soil and roots from old or modern wheat accessions, respectively, and were found all to be significantly greater than 0 (one-tailed t-test). Within each group, 96% of the bulk soil communities, 58% of the old and 33% of the modern communities had NRI values greater than 2. The average NRI values of the three groups were significantly different from each other (Figure S6).

The immigration rates for the bacterial and fungal communities were estimated from the model (Ntm value used to fit the model with the observations): 4.5% and 5% of the deaths happening in bacterial communities from modern and old accessions were replaced by OTUs from the source. Lower values were observed for the fungal communities: 0.9% and 1.4% for the modern and old community,
respectively. Among the old cultivars, we distinguished two further subsets: *T. aestivum* landraces including *Tas*, and *T. monococcum* and *T. turgidum* lines. The migration rates for these two subsets of communities were 6.2% and 3.8%, respectively. Results suggest that among the old accessions *Tm* and *Tt* lines created a more isolated environment, where fewer deaths happening within the local community were replaced by immigrants from the source community (soil), compared to the environment created by the *T. aestivum* landraces.

**Discussion**

The goal of this study was to identify potential ‘domestication effects’ by comparing root microbial communities in three distinct genetic groups of Triticum depicting two major events in domestication history: modern cultivars, landraces and ancestors of wheat. Breeding for higher yields in high-input agriculture may inadvertently have led to depletion of beneficial microbial taxa associated with plant roots. For example, Valente et al. [22] found that ancient wheat varieties were more capable of interacting with beneficial plant growth-promoting rhizobacteria. Generally, the magnitude of genotype effects on microbial communities is significant but small [3, 5, 26]. While a number of studies have investigated genotype effects in modern crop plants, fewer studies have specifically investigated domestication effects by comparing modern crops and the ancestors of crops [8, 9, 11, 12, 27]. In the present study, we demonstrated effects of domestication on root-associated microbial diversity, and we demonstrated that bacterial and fungal communities differed in a transect of evolution between the different groups of wheat: modern cultivars, landraces (including *T. triticum* ssp. *spelta*) and wheat ancestors (*T. turgidum* ssp. *dicoccum*, *T. monococcum* ssp. *monococcum* and *T. monococcum* ssp. *aegilopoides*). We were able to identify enriched bacterial and fungal taxa in old and modern wheat accessions, and we demonstrated that modern cultivars followed a neutral community assembly model to a higher degree than old lines, suggesting that the old lines posed a stronger selection on microbial community assembly.

Bacterial and fungal alpha diversity in bulk soil was higher than in root-associated communities, as has also been observed in numerous other studies, e.g. RL Berendsen, CMJ Pieterse and PAHM Bakker [1]. Both bacterial and fungal alpha diversity in root samples were generally lowest at BBCH10 and then dramatically increased and approached similar diversity levels at BBCH21. Similarly, Edwards et al. [28] demonstrated rapid colonization of rice roots even after a few days following transplantation of seedlings to soils, and succession of communities that stabilized after 2 weeks of plant growth. Overall, our results show that there is a rapid initial microbial colonization phase from bulk soil to the root, and we speculate that this period is of high importance in shaping community structures in plant roots, also at the later stages of growth.

Bacterial colonization of roots of modern cultivars occurred faster than in wheat ancestors (*Tm* and *Tt*), indicating that wheat ancestors subject stronger selection on bacteria at the initial growth stages. On the other hand, the lowest diversity of fungal communities was observed in modern cultivars at BBCH10. We speculate that one reason for this difference between bacterial and fungal communities could be that focus in plant breeding has been on fungal pathogens and that the plant immune system has been tuned
to prevent fungal invasion. Thus, modern cultivars may be more selective towards fungi from the earliest stages of root development, inhibiting rapid stochastic colonization. The highest diversities of both bacteria and fungi were observed in the wheat landraces, as was also observed in other studies [29, 30]. Landraces have been adjusted to agricultural settings during domestication, whereas wheat ancestors Tm and Tt have been harvested from native soils and are thus not adjusted to the microbial communities found in agricultural soils. At the same time, landraces were domesticated in low-input soils in opposite to modern cultivars that have been bred in high-input soils, promoting lower dependency on microbial interactions in the root.

Both fungal and bacterial communities differentiated according to bulk soil and wheat type (modern, landrace + Tas or ancient (Tm and Tt)) in NMDS plots. Bulk soil communities separated from root communities over time, with the most remarkable separation observed at BBCH21. We observed a stronger selection of bacterial and fungal communities at the later growth stages based on the more distinct clustering of samples at BBCH21. Landraces and ancient lines of wheat clustered more distinctly at later stages, which could indicate a stronger selection during the colonization of ancient lines, as was also observed by Özkurt et al. [31]. They studied endophytes in wheat and suggested that stochastic processes are of higher importance in modern wheat lines compared to ancient lines. They further speculated that domestication could have entailed less selective constraints on plant traits that contribute to microbial assembly. Bacterial communities clustered more distinctively at BBCH10 and BBCH12 compared to fungal communities suggesting a stronger selection of bacteria at the initial stages of colonization. Fungal and bacterial communities in the ancient wheat accessions clearly separated from modern cultivars already at BBCH10, showing that effects of domestication are manifested already at the initial growth stages of the wheat seedlings. Also of notice, both bacterial and fungal communities in modern cultivars separated from the communities in ancient and landrace accessions.

A number of bacterial and fungal taxa were enriched in old or modern accessions. We noted that more members of Acidobacteria and Actinobacteria were enriched in the old accessions, while more OTUs from Candidatus Saccharibacteria, Verrucomicrobia and Firmicutes were enriched in the modern cultivars. Interestingly, most of these OTUs were predominantly enriched in BBCH21. Acidobacteria, Verrucobacteria and Actinobacteria are typically found in high abundance in the soil and rhizosphere, and plant roots selectively enrich for certain groups of Acidobacteria and Verrucobacteria [32]. Actinobacteria and Firmicutes are associated with soil suppressiveness [33] and may thus be recruited as antagonists against soil-borne pathogens. At a lower taxonomic level, Rhizobiaceae, Flavobacteriaceae and Pedobacter were all enriched in modern cultivars. Interestingly, these taxa have all been suggested to have plant beneficial properties and to be enriched in the roots or rhizosphere of cereals [34–36]. Further corroborating this, C Zachow, H Muller, R Tilcher and G Berg [9] found a higher proportion of isolates that were active against phytopathogens in domesticated sugar beet in a comparative study of microorganisms of cultivated sugar beet and its wild ancestor Beta vulgaris ssp. maritima. Likewise, Bulgarelli et al. [11] found higher amounts of Flavobacteriaceae in the roots and rhizosphere of a modern barley cultivar compared to the ancestor Hordeum vulgare ssp. spontaneum. We observed enrichment of several known fungal wheat pathogens such as Fusarium, Neoascoscytha and Microdochium in the modern cultivars.
Soil-borne Fusarium was also enriched in modern cultivars of sunflower compared to less domesticated sunflower [8]. Whether enrichment of certain putative fungal pathogens in modern cultivars represents a general trend is unknown, but it is tempting to speculate that domesticated wheat may have lost the ability to prevent certain soil-borne pathogens, as those pathogens are not important in modern agricultural systems including efficient crop rotations.

Our results suggested that stochastic factors of birth, death and immigration from the bulk soil are largely driving the composition of both bacterial and fungal communities associated with wheat roots. However, there were still some discrepancies between the observed OTU frequencies and frequencies predicted by the NCM, suggesting that deterministic processes also play a role in shaping the local communities. In both bacterial and fungal root-associated communities, a percentage of the OTUs exhibited significant deviation from the NCM with an observed frequency either higher or lower than predicted by the model. Furthermore, ancient wheat lines created an environment where deterministic processes played a more predominant role in shaping the community compared to modern cultivars. Indeed, in modern cultivars communities, a higher percentage of OTUs was found to follow the NCM prediction. This finding is corroborated by the NMDS plots that showed a stronger clustering of communities belonging to the old wheat accessions, and by the observations of Özkurt et al. [31] who found higher stochasticity in the assembly of microbial communities in modern wheat. Strong phylogenetic clustering was observed in several microbial communities associated with roots. In wild oat, the core microbiome showed strong phylogenetic clustering, indicating that selective processes were dominating the community assembly [37]. We observed strongest phylogenetic clustering in bacterial communities associated to the ancient accessions (higher values of NRI), suggesting that environmental filtering is strongest in the root environment of ancient lines, promoting the co-occurrence of OTUs phylogenetically closer to each other, and thus sharing traits essential for their survival in such environments. Surprisingly, we observed a high phylogenetic clustering in the bulk soil. Microbial communities in fertile soils have been shown to have higher phylogenetic clustering than nutrient-depleted soils [38], promoting microbial taxa, such as Proteobacteria that prefer high resource availability [39]. The migration rates for landraces + Tas were much higher compared to the ancient lines Tm and Tt (6.2% vs. 3.8%). This suggests that Tm and Tt lines created a more isolated environment from the soil compared to the environment created by landraces. With a lower immigration rate, we also expected lower alpha diversity, and indeed, we observed that Tm and Tt had the lowest alpha diversity. The high diversity of the soil was lost in the local root communities and the more the local communities were isolated (low immigration rates) the less diversity we observed [40].

The results of our study suggest that changes have occurred during the course of domestication in the ability of wheat to assemble fungal and bacterial root communities. Whether these changes are a result of the selection for certain wheat traits under high nutrient availability remains an open question. It also remains to be elucidated whether such selection has led to changes in the ability of wheat to cope with environmental stresses, to take up nutrients from the soil more efficiently and to avoid pathogens in the root more effectively. However, indications of less dependency in modern wheat on microbial interactions have emerged. As an example, wheat dependence of mycorrhizal interactions was higher in landraces
compared to ancestors and modern lines, and responsiveness to mycorrhizal colonization was lower in modern cultivars [41].

**Conclusion**

In conclusion, we have demonstrated remarkable domestication effects in wheat. We have shown that microbial diversity in wheat roots grown in agricultural soils is highest in landraces domesticated in low-input agricultural systems. We also demonstrated that the first period after emergence of seedlings is crucial for microbial colonization of roots. Several taxa were enriched exclusively in modern or old wheat lines. Communities in both modern and old wheat lines followed the predictions of the neutral community assembly model, but modern cultivars exhibited stronger correlation with the model prediction, indicating a stronger selection for microbial taxa in old cultivars.

**Materials And Methods**

*Soil and plant sampling*

Accessions of old wheat lines were provided by NordGen, Alnarp, Sweden (https://www.nordgen.org/en/). The seeds of 5 accessions of wheat ancestors (*T. turgidum* ssp. *dicoccum* (*Tt*), *T. monococcum* ssp. *monococcum* (*Tm*) and *T. monococcum* ssp. *aegilopoides* (*Tm*)), 5 accessions of landraces (*T. aestivum* ssp *aestivum*), one accession of *T. aestivum* ssp. *spelta* (*Tas*) and 4 modern wheat cultivars (*T. aestivum* ssp *aestivum*) were sown on a sandy clay loam soil in the autumn 2017 in 50 cm rows at Aarhus University, Flakkebjerg, Denmark (55.322631°N, 11.394336°E) (Table 1). For data analysis, all accessions were categorized as modern cultivars, landraces, *Tas* or wheat ancestors, *Tt* and *Tm* (Table 1).

For each cultivar, root samples including tightly attached soil, were collected destructively in quadruplicates at BBCH10 (emergence of first leaf), BBCH12 (first leaf fully unfolded) and BBCH21 (first side shoot visible) [42] during the autumn 2017 starting from October 31 (Table S1). Each root sample was a composite of four plants, which were randomly uprooted and carefully freed from the soil loosely attached soil.

Bulk soil samples were collected from the plough layer (0-20 cm) in between crop rows at each sampling point. In each sampling line (space between two crop rows) 10 soil sub-samples were collected, thoroughly mixed and from those two replicate composite bulk soil samples were collected (n=10). All samples were frozen at -80°C until further processing for DNA extraction.

*Sample homogenization and DNA extraction*

Soil and root samples were freeze-dried for 48h at -111°C and 0.0026 Pa using a CoolSafe CS110-4 Pro freeze dryer (LaboGene, Denmark) and finely crushed using zirconium oxide grinding balls (10mm) and a Fast & Fluid Management shaker-SK350 (Fast & Fluid, Sassenheim Netherlands). DNA was extracted from homogenized soil and root samples using the NucleoSpin® DNA 96 Soil Kit (Macherey-Nagel GmbH
& Co KG, Düren, Germany) adapted to a Biomek® FCP Laboratory Automation Workstation (Beckman Coulter™, CA, USA). Four negative controls (500µL of PCR graded water) were included per extraction plate and used for amplicon library preparation together with the samples. Concentration of extracted DNA was measured using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, MA, USA) with a Qubit dsDNA HS Assay Kit (range 0.2-100ng; Invitrogen, Life Technologies). DNA was then stored at -20°C.

**Amplicon library preparation and sequencing**

Extracted DNA was used to produce bacterial and fungal amplicon libraries (Table S2). The fungal ITS2 region was amplified using fITS7 and ITS4 primers (Ihrmark *et al.*, 2012) and the bacterial 16S rRNA V3-V4 region was targeted using primers S-DBact-0341-b-S-17/S-D-Bact-0785-a-A-21 [43]. PCR was performed in a reaction mixture of 25 µl consisting of 1 x PCR reaction buffer, 1.5 mM MgCl2, 0.2 mM dNTPs, 1 µM of each primer, 1 U of GoTaq Flexi polymerase (Promega Corporation, Madison, USA) and 1µl of DNA template. PCR was conducted in a GeneAmp PCR System 9700 thermal cycler (Thermo Fisher Scientific) using 94 °C for 5 min, followed by 25 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. For fungal amplicon library preparation, an annealing temperature of 57 °C was used as recommended [44]. For dual indexing, primers including indexing tags were used in a PCR for 10 cycles, with the thermal cycler program as described above. In addition to dual indexing, internal barcodes of varying length were added to the forward primer for combining samples within each index combination as described earlier [45, 46]. After PCR, amplicon size was confirmed by visualization in a 1.5% agarose gel using SYBR staining, PCR products were pooled, precipitated and re-eluted as described earlier [47]. In order to remove primers and shorter reads, the pooled DNA was separated on a 1.5 % agarose gel and amplicons of the expected size (300–450 bp) were extracted using a QIAquick Gel Extraction Kit (Qiagen, Copenhagen, Denmark). The DNA concentration of the amplicon library was evaluated using a Qubit® Flurometer (Thermo Fisher Scientific). Amplicon libraries were shipped to Eurofins MWG (Ebersberg, Germany) for sequencing on an Illumina MiSeq platform using a dual indexing strategy.

**Amplicon sequencing data analysis**

Quality processing and bioinformatics of the reads obtained from Illumina MiSeq were analysed as described earlier [48]. Briefly, paired end reads were filtered with Phred quality scores > 30 and merged with an overlapping minimum read length of 30 base pairs using VSEARCH v. 2.6 [49]. Forward and reverse primers, as well as internal barcodes were trimmed and reads with less than 200 base pairs were excluded. Before clustering at 97% similarity, all reads were dereplicated and screened using VSEARCH. Taxonomy assignments for the clustered operational taxonomic units (OTUs) was done using SILVA v. 132 for bacteria and UNITE v. 8.0 for fungi in QIIME v 1.9, using assign_taxonomy.py [50, 51]. OTUs unassigned at kingdom level or assigned as chloroplast or mitochondrial sequences were removed from the datasets. Furthermore, samples with less than 2000 reads were excluded from downstream analysis.
Bacterial and fungal annotation tables were analyzed in R v.3.5.2 (R Core Team, 2017), using the RStudio development environment (RStudio Team) and making use of the ‘phyloseq’ package [52]. The OTU tables were transformed, by either rarefaction or relative abundance, before executing diversity-based calculations. Alpha diversity was estimated using observed OTU richness and Shannon diversity measures, using the mean value from rarified OTU tables generated 100 times at a sampling depth of 2000 reads per sample. Significant differences between diversity indices were determined using the Kruskal-Wallis rank sum test. Bray-Curtis distance matrices were used for beta diversity analysis on OTU level representing > 2000 reads per sample and transformed to relative abundance. Variance partitioning and significance for experimental factors were detected by PERMANOVA (R package: ‘vegan’). Dissimilarity between samples (based on Bray-Curtis distances) was visualized using NMDS plots.

To study the importance of selection versus neutral processes in community assembly, a modified version [53] of the model described by Sloan et al. [25] was applied. This neutral community assembly model (NCM) predicts the frequencies with which taxa should occur in target communities based on their abundance in the source community (also referred to as metacommunity). In short, the model predicts that abundant taxa in the source community will be observed more frequently in the target communities due to increased immigration opportunities, while rare taxa will more likely be lost in the target community due to ecological drift. The model fit is determined by the coupled parameter \( N_t \times m \); where \( N_t \) is the size of the community (the average number of reads was considered as an estimation of the community size); \( m \) is the migration rate. The migration rate is the estimated probability that the random loss of an individual in the local community will be replaced by immigration from the source community. The fitting of the model parameters was performed in R where binomial proportion 95% confidence intervals (based on the Wilson method) around the model predictions were calculated using the ‘HMisc’ package [54]. In this study, root samples were considered as local communities, while the composition of the source community was inferred by averaging the composition of the bulk soil samples. Among the local communities, two separate groups were identified: roots communities associated with the modern opposed to old cultivars; these two analyses were performed for the fungal and bacterial communities separately. OTUs falling between the 95% confidence interval were considered to be present as a result of neutral dynamics of birth, death and immigration from the source (bulk soil); OTUs falling outside the confidence interval (and with the ratio between observed frequency and predicted frequency greater than 1.5 or lower than 0.5) were found with frequencies disproportionally higher or lower than predicted by the model based on their abundances in the soil and therefore considered to be present in the communities as a result of deterministic factors. Pearson Correlation between the frequency of the sequence variants estimated by the model and that of observed was used to determine the statistical significance of model fit. The analysis was performed considering only OTUs shared by the bulk soil community (source) and the root community (target). In this case, all sequence variants that were detected in the root, but not in the bulk soil, and vice versa were excluded, as it is routinely done [53, 55, 56].
Net relatedness index (NRI), an index of community phylogenetic structure, was used to study the phylogenetic community composition within each sample [57, 58]. NRI quantifies the grade of phylogenetic clustering of a given community compared to a stochastically assembled community composed of taxa from the regional pool. The NRI was calculated by comparing the observed mean pairwise phylogenetic distance (MPD) of OTUs in each sample with the MPD of a null distribution, obtained by randomising the OTUs of the regional pool and their relative abundances across the tips of phylogeny. Values of NRI greater than +2 are typical of communities where taxa appear more closely related than expected by chance (phylogenetic clustering); while values of NRI less than -2 indicate the presence of taxa more distantly related (phylogenetic over-dispersion) [59]. The function ‘ses.mpd’ in the package ‘picante’ was used to calculate the index, employing the first 6000 most abundant OTUs and using ‘taxa.labels’ as the reference null model [60].

Average NRI values were calculated for bacterial communities from bulk soil (SOURCE); from roots of old accessions (OLD) and modern cultivars (MODERN). This analysis was employed to study the relative contribution of stochastic and deterministic factors on the phylogenetic assembly in the three different environments. Parametric and non-parametric analyses were performed to ask whether the average NRI values were statistically different from zero and between the three environments.

Declarations

Acknowledgements

Novozymes A/S would like to acknowledge Birgitte Gjerde Bennedsen for help with the lab work. Mathilde Schiøtt Dige, Aarhus University is likewise acknowledged for her lab work.

Author’s contributions

MN and IMN designed the experiments. IMN collected the samples. IMN and RS performed the experiments. MKG, RS and MV performed the bioinformatics. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Funding

Funding for R.K and M.N. was provided by Independent Research Fund Denmark, grant no. 6111-00065B. M.V. was supported by the Royal Academy of Engineering under the Research Fellowship scheme.

Availability of data and materials

All amplicon sequencing files from this study were deposited in the NCBI sequence read archive under the accession number PRJNA611715.

Ethics approval and consent to participate

Not applicable
Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests

References


**Tables**

*Table 1.* Triticum accessions used.
<table>
<thead>
<tr>
<th>Name</th>
<th>Accession</th>
<th>Genome</th>
<th>Category</th>
<th>Scientific name</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metalla</td>
<td>NGB22783</td>
<td>AA</td>
<td></td>
<td><em>T. monococcum</em> (<em>Tm</em>) <em>T. monococcum</em> ssp <em>monococcum</em></td>
<td>Cultivated Einkorn</td>
</tr>
<tr>
<td>Einkorn</td>
<td>NGB22756</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRI 1774</td>
<td>NGB10931</td>
<td></td>
<td></td>
<td><em>T. monococcum</em> ssp <em>aegilopoides</em></td>
<td>Wild Einkorn</td>
</tr>
<tr>
<td>KVL 2379</td>
<td>NGB4802</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emmer</td>
<td>NGB8190</td>
<td>BBA</td>
<td><em>T. turgidum</em>(Ti)</td>
<td><em>T. turgidum</em> ssp <em>dico</em>con</td>
<td>Emmer</td>
</tr>
<tr>
<td>Haillets</td>
<td>NGB8967</td>
<td>BBAADD</td>
<td>Landrace</td>
<td><em>T. aestivum</em> ssp <em>aestivum</em></td>
<td>Known in UK before 1900</td>
</tr>
<tr>
<td>ALS</td>
<td>NGB4770</td>
<td></td>
<td></td>
<td></td>
<td>Known in Denmark in the 1940s</td>
</tr>
<tr>
<td>Kolbe</td>
<td>NGB7032</td>
<td></td>
<td></td>
<td></td>
<td>Known before 1900</td>
</tr>
<tr>
<td>Lys Østpreussisk Hvede</td>
<td>NGB8979</td>
<td></td>
<td></td>
<td></td>
<td>Known before 1900</td>
</tr>
<tr>
<td>Square Head II</td>
<td>NGB5147</td>
<td></td>
<td></td>
<td></td>
<td>Known from the 1870s</td>
</tr>
<tr>
<td>Oberkulmer (selection)</td>
<td>NGB22516</td>
<td></td>
<td><em>Spelt</em> <em>Tas</em></td>
<td><em>T. aestivum</em> ssp <em>spelta</em></td>
<td>Spelt</td>
</tr>
<tr>
<td>KWS Desanto</td>
<td>N/A</td>
<td></td>
<td>Modern cultivars</td>
<td><em>T. aestivum</em> ssp <em>aestivum</em></td>
<td>Susceptible to <em>Blumeria</em> <em>graminis</em> and <em>Zymoseptoria</em> <em>tritici</em></td>
</tr>
<tr>
<td>Substance</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td>Resistant to <em>Blumeria</em> <em>graminis</em> and <em>Zymoseptoria</em> <em>tritici</em>, susceptible to <em>Puccinia</em> <em>striiformis</em></td>
</tr>
<tr>
<td>Sheriff</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td>Resistant to <em>Blumeria</em> <em>graminis</em>,</td>
</tr>
</tbody>
</table>
**Table 2.** Permutation analysis of variance (PERMANOVA) of root-associated bacterial and fungal communities in the different wheat genotypes (modern cultivars, Landraces + *Tas* and ancient lines). Adonis tests were based on Bray-Curtis distance matrices using 1000 permutations. After the initial analysis of the total dataset, bulk soil samples were removed from the dataset, and the remaining data was split into growth stage.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Factor</th>
<th>Df</th>
<th>Bacteria ($R^2$)</th>
<th>Fungi ($R^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Bulk soil vs root</td>
<td>1</td>
<td>0.13***</td>
<td>0.06***</td>
</tr>
<tr>
<td>Roots</td>
<td>Triticum genetic group</td>
<td>3</td>
<td>0.11***</td>
<td>0.08***</td>
</tr>
<tr>
<td></td>
<td>Growth stage (BBCH)</td>
<td>2</td>
<td>0.06***</td>
<td>0.05***</td>
</tr>
<tr>
<td></td>
<td>Triticum genetic group * BBCH</td>
<td>6</td>
<td>0.12***</td>
<td>0.05***</td>
</tr>
<tr>
<td>Roots BBCH10</td>
<td>Triticum genetic group</td>
<td>3</td>
<td>0.23***</td>
<td>0.13***</td>
</tr>
<tr>
<td>Roots BBCH12</td>
<td>Triticum genetic group</td>
<td>3</td>
<td>0.17***</td>
<td>0.17***</td>
</tr>
<tr>
<td>Roots BBCH21</td>
<td>Triticum genetic group</td>
<td>3</td>
<td>0.34***</td>
<td>0.11**</td>
</tr>
</tbody>
</table>

**Table 3.** Number of OTUs present in the target communities (roots), in the source community (bulk soil) and the OTUs shared between the communities and retained for the NCM analyses. Among the shared OTUs, the number of OTUs that are following neutrality, are advantaged or are disadvantaged, are shown.
<table>
<thead>
<tr>
<th>Shared OTUs:</th>
<th>Bacteria</th>
<th>Fungi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantaged OTUs</td>
<td>794</td>
<td>13%</td>
</tr>
<tr>
<td>Disadvantaged OTUs</td>
<td>685</td>
<td>12%</td>
</tr>
<tr>
<td>Neutral OTUs</td>
<td>4462</td>
<td>75%</td>
</tr>
</tbody>
</table>

**Figures**
Figure 1

Diversity indices (richness and Shannon Index) calculated at OTU level for a) bacterial communities rarefied to 3000 reads and b) fungal communities rarefied to 5000 reads. Error bars represent the standard deviation of average values.
Figure 2

Nonmetric multidimensional scaling (NMDS) ordination of (A) bacterial and (B) fungal community composition in bulk soil and roots of modern, landrace + Tas, and Tm and Tt wheat accession at each sampled growth stage based on Bray-Curtis distances.
Figure 3

Bacterial (A) and fungal (B) phyla differentially enriched in the roots of old cultivars (log2Fold change <0) or modern cultivars (log2Fold change >0). “Old accessions” depicted in this figure groups the wheat ancestors Tm and Tt, landraces and Tas. “Modern cultivars” groups the commercial cultivars KWS Desanto, Substance, Sheriff and Torp.
**Figure 4**

Fit of the predicted neutral model (NCM) to observed bacterial (top panels) and fungal community (bottom panels) assembly at OTU level. The predicted occurrence frequencies for OTUs shared between bulk soil and roots in (A, C) modern and (B, D) old wheat accessions and the predicted occurrence frequencies. OTUs that occur significantly more frequently than predicted by the model are shown in green, while those that occur less frequently than predicted are shown in red. Dashed lines represent 95% confidence intervals around the model prediction (black line), $m =$ immigration rate, $\rho =$Pearson correlation.

**Supplementary Files**

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

- Supplementaldata.docx