Global diversity of putative functional viral auxiliary metabolic genes and their relationship with soil organic carbon stocks

Dominik Merges*

*Corresponding author:
Dominik Merges
Mergesd01@gmail.com

Running title: Functional AMG diversity in soils

Key words: bacteria, bacteriophages, ecosystem functioning, ecosystem processes, nutrient cycling, carbon, nitrogen, species interactions

Abstract

Global biodiversity decline has prompted extensive research and debate on how biodiversity loss affects ecosystem functioning. Microbial taxonomic and functional diversity loss can impair critical processes such as plant biomass degradation or nitrogen fixation. Biotic interactions, especially those involving microbes and their viruses, are increasingly recognized as drivers of such processes. Microbial viruses carry auxiliary metabolic genes (AMGs), functional genes acquired from bacterial hosts, that mediate processes like carbon and nitrogen cycling, positioning viruses as potential reservoirs of essential ecosystem functions. To investigate the putative contribution of viral AMGs to microbial functional diversity and ecosystem functioning in soils, the global dataset of soil viral genomes, comprising 616,935 uncultivated viral genomes and 38,508 unique viral operational taxonomic units from 2,953 soil metagenomes, was mined for carbon (C) and nitrogen (N) cycling genes. For C cycling, 301 viral genomes with AMGs from glycoside hydrolase (GH) families (n cellulase = 182; n hemicellulase = 119) were detected. The viruses containing cellulase and hemicellulase AMGs were distributed across most major ecosystem types, including agricultural, forest, grassland, and permafrost soils, but were absent from deserts, salt marshes, and shrublands. Forest ecosystems harbored the highest abundance of cellulase and hemicellulase AMGs. The ratio of cellulase to hemicellulase AMGs varied by ecosystem, showing near parity in forests but cellulase dominance in agricultural, grassland, and permafrost soils. Relating AMGs to ecosystem functioning, sites with cellulase and hemicellulase AMGs showed significantly lower soil organic carbon (SOC) content in agricultural soils, whereas no significant differences were observed in forest, grassland, or permafrost soils. These results suggest that viral AMGs may influence SOC dynamics, particularly in agricultural systems, and highlight the need for deeper investigation into their functional roles in soil ecosystems. This study underscores the importance of viral contributions to ecosystem functioning, emphasizing the potential of microbial viruses to serve as mutualistic agents in global biogeochemical cycles.

Global biodiversity decline has prompted extensive research and debate on how biodiversity loss affects ecosystem functioning (WWF 2022). Microbial taxonomic and functional diversity loss can impair critical processes such as plant biomass degradation or nitrogen fixation, and denitrification. The loss of microbial taxonomic and functional diversity can impair processes such as plant biomass decomposition (Wagg et al., 2019), nitrogen fixation, and denitrification (Castaño et al., 2023; Philippot et al., 2013). However, in certain instances, diversity loss appears to have little to no impact on these processes, highlighting the complexity of interactions within microbial communities (Wertz et al., 2006). These interactions may be dependent on attributes of communities differing in their structure and species composition that may result in different patterns of ecosystem functional response to diversity loss (Tilman et al., 1997; Philippot et al., 2013). Currently, it remains uncertain which specific community attributes are most relevant to an ecosystem's functional response to the loss of diversity.

Among the community attributes relevant to ecosystem processes and functions, biotic interactions, particularly those involving microbes and their viruses, are increasingly recognized as relevant (Braga et al., 2018). While significant attention has been given to viruses' antagonistic roles, their potential mutualistic effects on ecosystem functioning remain relatively unexplored. Recent findings indicate that microbial viruses can carry auxiliary metabolic genes (AMGs), functional genes acquired from bacterial hosts, that play critical roles in processes like carbon and nitrogen cycling, stress response, and energy production, positioning viruses as potential reservoirs of essential ecosystem functions (Roux et al., 2016).

Although most AMG-related evidence originates from marine systems (Roux et al., 2016), soil studies have also demonstrated virus-encoded AMGs supporting bacterial and archaeal functions, including pesticide degradation, carbon and methane cycling, and nitrification (Emerson et al., 2018; Zheng et al., 2022). Moreover, specific functional AMGs have been experimentally verified for roles in plant biomass decomposition (Emerson et al., 2018) and organochlorine pesticide degradation (Zheng et al., 2022). These findings suggest that microbial viruses may function as mutualists, equipping microbial communities with additional functional genes and thereby significantly influencing ecosystem functions (Jurburg et al., 2023).

The aim of this study was to investigate 1) the prevalence of putative functional AMGs across different ecosystem types and 2) their potential contribution to microbial functional diversity and ecosystem functioning in soils.

To quantify the prevalence of viral AMGs in soils, functionally annotated viral genomes from the global dataset of soil viral genomes (Graham et al. 2024) were screened, to collect viral genomes with AMGs linked to ecosystem functioning. The global dataset of soil viral genomes contained 616935 uncultivated viral genomes and 38508 unique viral operational taxonomic units assembled from 2953 previously sequenced soil metagenomes (Graham et al. 2024). From this dataset, a database of viral genomes with functional AMGs, relevant for C (i.e. degradation of plant matter) and N (nitrogen fixation) cycles (Figure 1a) was build, by identifying AMGs linked to C and N cycles, leveraging functional annotations from CAZy database (Drula et al. 2022 and the KEGG database modules to phenotype annotations (Kanehisa et al. 2016). For C cycling, the viral genomes were screened for AMGs associated with the degradation of plant biomass, i.e. microbial GH families targeting plant cell wall degradation, specifically for cellulose (CAZyme identifier: GH5, GH6, GH7, GH9, GH16, GH17, GH44, GH45, GH48, GH74; Auer et al. 2023) and hemicellulose (CAZyme identifier: GH8, GH10, GH11, GH12, GH26, GH28, GH30, GH53; Auer et al. 2023). In total, 301 viral genomes with AMGs from the GH families (n cellulase = 182; n hemicellulase = 119) were found (Fig. 1a). For N cycling, I found one viral genome with a nifH AMG. The NifH gene encodes for the Nitrogenase iron enzyme responsible for nitrogen fixation of nitrogen-fixing bacteria and archaea (nifH, KEGG ortholog K02588; Fig 1a).

The cellulase and hemicellulase AMGs were present within most major ecosystem types (i.e. agricultural, forests, grasslands, permafrost), but absent in deserts, salt marshes and shrublands (Fig. 1b). Cellulase and hemicellulase AMGs were most common in forest ecosystems, which typically have rather stable environmental conditions and a highly diverse plant biomass, including leaves, wood and understorey vegetation which contain a mix of cellulose and hemicellulose, and accordingly, both cellulase and hemicellulase AMGs are abundant. The two degradation types (i.e. cellulase and hemicellulase) show different ratios across ecosystems, with an almost balanced ratio in forests, but dominance of cellulase over hemicellulose in agricultural, grassland and permafrost ecosystems (Fig. 1c). Cellulase dominance may be attributed to several factors, such as that microorganisms may prioritize cellulase production 1) based on the lower diversity of plant biomass in agricultural, grassland, and permafrost soils and 2) under less stable environmental conditions (e.g., tillage, moisture variability, and temperature extremes), as cellulose provides a more recalcitrant and consisted carbon source compared to hemicellulose.

Relating the presence of cellulase and hemicellulase AMGs to ecosystem functioning, a comparison of soil organic matter (SOC) content in sites lacking degradation genes (grey boxes) versus those containing degradation genes (blue boxes) reveals significantly lower SOC content in agricultural samples harboring putative viral degradation genes (t-test $p < 0.001$; Fig. 1d). This observation aligns with findings that agricultural practices accelerate SOC breakdown through microbial activity (Lal, 2004; Six et al., 2002). The presence of viral degradation genes may indicate active viral and microbial communities contributing to SOC decomposition, consistent with evidence linking viral auxiliary metabolic genes (AMGs) to ecosystem-level carbon cycling (Roux et al., 2016). In agricultural systems, intensive land-use practices (e.g., tillage, crop harvest) expose organic matter to microbial degradation, potentially enhancing SOC breakdown by active viral and microbial communities (Paustian et al., 1997). Conversely, the stable microclimate and continuous organic matter input in forests, low temperatures and frozen soils in permafrost, as well as extensive root biomass and well-structured soils in grasslands, may contribute to maintaining stable SOC levels regardless of the presence of viral AMGs (Jobbágy & Jackson, 2000; Schuur et al., 2015; Jackson et al., 1996).

The observed association of putative viral AMG presence on SOC in agricultural systems underscores the need for a more comprehensive evaluation of their ecological roles. Controlled experiments offer a promising approach to elucidate the contribution of viral AMGs to microbial-mediated SOC degradation. Such investigations will advance our understanding of the interplay between viral and microbial communities in shaping ecosystem carbon dynamics.

References

Burke, C., P. Steinberg, D. Rusch, S. Kjelleberg, T. Thomas, Bacterial community assembly based on functional genes rather than species, Proc. Natl. Acad. Sci. U.S.A. 108 (34) 14288-14293.

Braga, L. P., Soucy, S. M., Amgarten, D. E., Da Silva, A. M., & Setubal, J. C. (2018). Bacterial diversification in the light of the interactions with phages: the genetic symbionts and their role in ecological speciation. Frontiers in Ecology and Evolution, 6, 6.

Castaño, C., Hallin, S., Egelkraut, D., Lindahl, B.D., Olofsson, J. and Clemmensen, K.E. (2023), Contrasting plant–soil–microbial feedbacks stabilize vegetation types and uncouple topsoil C and N stocks across a subarctic–alpine landscape. New Phytol, 238: 2621-2633.

Emerson, J. B., Roux, S., Brum, J. R., Bolduc, B., Woodcroft, B. J., Jang, H. B., ... & Sullivan, M. B. (2018). Host-linked soil viral ecology along a permafrost thaw gradient. Nature microbiology, 3(8), 870-880.

Graham, E.B., Camargo, A.P., Wu, R. et al. A global atlas of soil viruses reveals unexplored biodiversity and potential biogeochemical impacts. Nat Microbiol 9, 1873–1883 (2024).

Grime, J. Competitive Exclusion in Herbaceous Vegetation. Nature 242, 344–347 (1973.

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., & Schulze, E. D. (1996). A global analysis of root distributions for terrestrial biomes. Oecologia, 108(3), 389–411. Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological Applications, 10(2), 423–436. Jurburg SD, Hom EFY, Chatzinotas A (2023) Beyond pathogenesis: Detecting the full spectrum of ecological interactions in the virosphere. PLOS Biology 21(5): e3002109.

Kuzyakov, Y., & Mason-Jones, K. (2018). Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. Soil Biology and Biochemistry, 127, 305-317.

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. Science, 304(5677), 1623–1627. McNaughton S J. Biodiversity and function of grazing ecosystems. In: Schulze E D, Mooney H A, editors. Biodiversity and ecosystem functioning. New York, N.Y: Springer Verlag; 1993. pp. 361–384.

McNaughton S J. Biodiversity and function of grazing ecosystems. In: Schulze E D, Mooney H A, editors. Biodiversity and ecosystem functioning. New York, N.Y: Springer Verlag; 1993. pp. 361–384.

Merges, D., Dal Grande, F., Greve, C., Otte, J., & Schmitt, I. (2021). Virus diversity in metagenomes of a lichen symbiosis (Umbilicaria phaea): complete viral genomes, putative hosts and elevational distributions. Environmental Microbiology, 23(11), 6637-6650.

Merges, D., Schmidt, A., Schmitt, I., Neuschulz, E. L., Dal Grande, F., & Bálint, M. (2023). Metatranscriptomics reveals contrasting effects of elevation on the activity of bacteria and bacterial viruses in soil. Molecular Ecology, 32(23), 6552-6563.

Naeem, S., Thompson, L. J., Lawler, S. P., Lawton, J. H., & Woodfin, R. M. (1994). Declining biodiversity can alter the performance of ecosystems. Nature, 368(6473), 734-737. Naeem S. (1998). Species redundancy and ecosystem reliability. Conserv Biol 12: 39–45

Le Provost G., Thiele J., Zaitsev A. S., Manning P. (2021): Contrasting responses of above- and belowground diversity to multiple components of land-use intensity. Nature Communications 12:3918.

Roux, S. et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. Nature 537, 689–693 (2016).Tilman, D., Reich, P. B. & Knops, J. M. H. Biodiversity and ecosystem stability in a decade-long grassland experiment. Nature 441, 629–632 (2006).

Tilman D, Lehman C, Thomson K. (1997). Plant-diversity and ecosystem productivity: theoretical considerations. Proc Natl Acad Sci USA 94: 1857–1861.

Paustian, K., Collins, H. P., & Paul, E. A. (1996). Management controls on soil carbon. In E. A. Paul et al. (Eds.), Soil organic matter in temperate agroecosystems (pp. 15–49). CRC Press.

Philippot, L. et al. Loss in microbial diversity affects nitrogen cycling in soil. ISME J. 7, 1609–1619 (2013).

Santos-Medellín, C., Estera-Molina, K., Yuan, M., Pett-Ridge, J., Firestone, M. K., & Emerson, J. B. (2022). Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands. Proceedings of the National Academy of Sciences, 119(45), e2209132119.

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. Nature, 520(7546), 171–179.

Six, J., Conant, R. T., Paul, E. A., & Paustian, K. (2002). Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. Plant and Soil, 241(2), 155–176.

ter Horst, A. M., Jane D Fudyma, Jacqueline L Sones, Joanne B Emerson, Dispersal, habitat filtering, and eco-evolutionary dynamics as drivers of local and global wetland viral biogeography, The ISME Journal, Volume 17, Issue 11, November 2023, Pages 2079–2089.

Trubl, G., Kimbrel, J.A., Liquet-Gonzalez, J. et al. Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. Microbiome 9, 208 (2021).

van der Plas, F. (2019). Biodiversity and ecosystem functioning in naturally assembled communities. Biological Reviews of the Cambridge Philosophical Society, 94(4), 1220–1245.

Wagg, C., Schlaeppi, K., Banerjee, S. et al. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. Nat Commun 10, 4841 (2019.

Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Freitag T et al. (2006). Maintenance of soil functioning following erosion of microbial diversity. Environ Microbiol 8: 2162–2169.

WWF. (2022). Living planet report 2022—building a nature-positive society. M. Grooten, D. Bignoli Juffe, & T. Petersen (Eds.). WWF.

Zheng, X., Jahn, M.T., Sun, M. et al. (2022) Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes. ISME J 16, 1397–1408

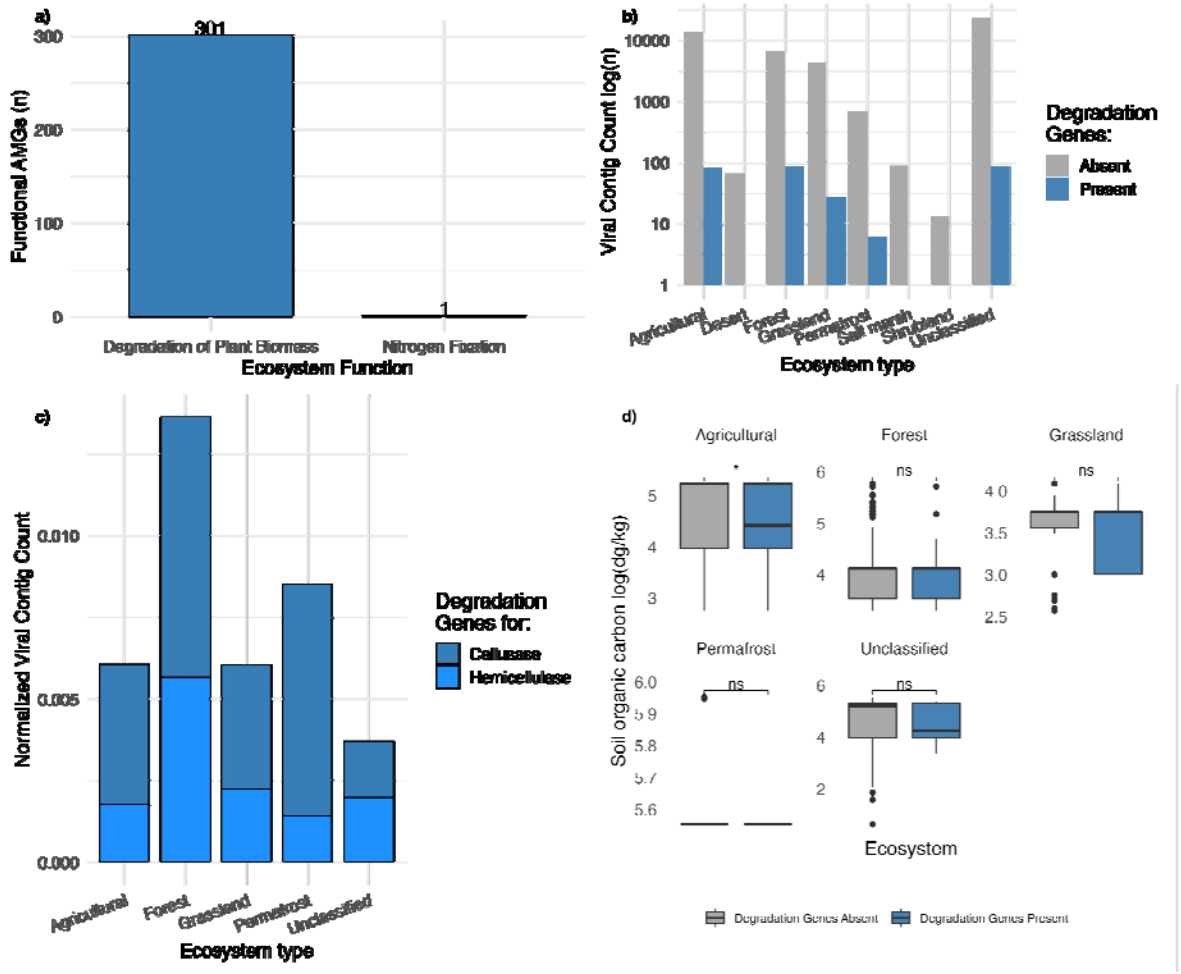

Figure 1: A) Number of found viral genomes containing AMGs for the degradation of plant biomass and nitrogen fixation. B) Prevalence of putative cellulase and hemicellulase AMGs. C) Putative cellulase and hemicellulase AMGs distributions differ across ecosystems D) Agricultural sites with cellulase and hemicellulase AMGs hold significantly less SOC than agricultural without. In forest, grassland and permafrost ecosystems sites with and without cellulase and hemicellulase AMGs do not differ in SOC content.