

Recurring perturbations limit the length of byproduct cross-feeding chains in digital communities

Johanna Orsholm

Examinator, Anna Eklöf
Tutor, György Barabás

Contents

1	Abstract	1
2	Introduction	1
3	Materials and methods	3
3.1	Avida software for digital evolution	3
3.2	Environmental configurations and the evolution of digital communities . .	3
3.3	Perturbation experiments	4
3.4	Phenotype and cross-feeding chain length	5
3.5	Lineage tracing	5
4	Results	5
5	Discussion	10
5.1	Conclusions	12
6	Societal and ethical considerations	13
7	Acknowledgments	13
	Appendix	17
A	Task order	17
B	Evolution of chain length in perturbed communities	19

1 Abstract

The human gut microbiome is important for health and development, and understanding its functioning and dynamics are of great medical importance. The microbiome food web is largely characterized by chains of byproduct cross-feeding (where metabolites of one organism are used as nutrients for another), yet a recent study have shown that the average length of the chains are considerably shorter than what metabolic capabilities of present species allow for. Here, I use evolving populations of digital organisms to investigate if recurring perturbations are a potential constraint of byproduct cross-feeding chains. I evolved digital populations in an environment unconstrained by energy loss between trophic levels and then exposed them to a period of recurring perturbations, where a fraction of the population was removed at 100 random points in time. Perturbations caused a substantial decrease in cross-feeding chain length, with increased frequency as perturbation intensity increased. In some communities, effects persisted after the perturbation period had ended. Tracking evolution of resource use during and after the perturbation period revealed that organisms descending from long-chained ancestors often evolved a shorter chain, suggesting that they adapted to perturbations by losing the ability to consume low-level resources. The evolutionary loss of resource consumption could explain the persisting effects on cross-feeding chains. Though my study suggests that perturbations can limit the length of byproduct cross-feeding chains, further studies are necessary to conclude if effects remain in environments with a more realistic energy transfer between trophic levels.

Keywords: Byproduct, Cross-feeding, Digital evolution, Gut microbiome

2 Introduction

The human gut microbiome is a complex ecological system of great significance for health and development (Cho and Blaser, 2012; Brestoff and Artis, 2013). Changes in its composition and function has been linked to multiple diseases, such as diabetes (Qin et al., 2012) and inflammatory bowel disease (Franzosa et al., 2019), suggesting that manipulation of the gut microbiome has potential for preventing or treating diseases. To make use of this potential, we must increase our understanding of the ecological and evolutionary processes that influence microbiome assembly and functioning.

The gut microbiome is largely characterized by cross-feeding interactions, where microbes consume metabolites produced by another individual (Sung et al., 2017). These cross-feeding interactions can be organized into trophic levels, characterized by metabolites being transferred from higher to lower levels (Wang et al., 2019). In other words, the microbiome food web largely consists of chains of cross-feeding interactions, where organism A creates a byproduct consumed by organism B, which in turn creates a byproduct

consumed by organism C, and so on. The average length of such chains in the human gut has been found to be four steps (Wang et al., 2019). However, mappings of metabolites and metabolic capabilities of present species, based on experimental evidence, reveal that considerably longer chains are theoretically possible (Sung et al., 2017). Thus, it is not yet known why these chains are so severely constrained, when they need not be in principle.

Here, I use evolving populations of digital organisms to investigate constraints to multi-level byproduct cross-feeding. Digital organisms are self-replicating computer programs that mutate and compete for their basic resource: computational power that enables execution of their genetic program. Meeting the conditions for natural selection (variation, differential fitness, and heritability; Lewontin, 1970), the system enables open-ended digital evolution. As opposed to mathematical models of evolutionary processes, digital evolution can be considered an instance of evolution, rather than a mere simulation (Pennock, 2007). While mathematical models always reflect the assumptions used to build them, digital evolution can result in unexpected traits and interactions (for some examples of surprising adaptations in experiments with digital evolution, see Lehman et al., 2018). Thus, by using a system for digital evolution, I free the experiments from any preconceptions of which mechanisms govern the results.

Avida is a widely used software platform for digital evolution, enabling detailed control of the environment, a clear separation of the effects of different factors, and a complete record of the course of evolution (Ofria and Wilke, 2004). Among other things, it has previously been used to study evolution of complex traits (Lenski et al., 2003), adaptive radiation (Cooper and Ofria, 2002a; Chow et al., 2004), and mutational robustness (Lenski et al., 1999; Elena et al., 2007). Avida consists of a virtual grid inhabited by digital organisms, each with a genetic program containing instructions for self-replication through copying (Ofria and Wilke, 2004). An imperfect copying process occasionally cause errors during replication, thus introducing mutations in the offspring's genome. Space-limitation in the virtual world cause an evolutionary pressure for rapid replication, and organisms effectively compete for computational power by consuming virtual resources. The environment can be configured to contain multiple resources, and to consume each resource, organisms must perform an associated computational task (Cooper and Ofria, 2002b). As a unique task is coupled to each resource, it is possible for organisms to specialize on different resources, thus enabling adaptive radiation. Configuring each resource to generate a byproduct when consumed, it is possible for cross-feeding interactions to evolve in the digital community.

In this work, I aim to investigate if recurring perturbations constrain the length of cross-feeding chains, using evolving digital populations. Using a digital study system, I will be able to separate the effects of perturbations from, for example, constraints imposed by energy loss between trophic levels. Further, Avida allows tracking of evolutionary adaptations, enabling analyses of evolutionary processes affecting community functioning in a perturbed environment.

3 Materials and methods

3.1 Avida software for digital evolution

All experiments were performed in version 2.14 of the Avida software platform for digital evolution.¹ Default settings were used unless otherwise indicated. In Avida, self-replicating computer programs – digital organisms – evolve by means of mutation, natural selection, and genetic drift. Organisms, also referred to as ‘Avidians’, each have a circular genome composed of programming instructions. By executing their genetic program, they are capable of self-replication through copying. An imperfect replication process can cause substitutions of one programming instruction for another, or ‘mutations’, in the offspring’s genome.

Avidians inhabit a virtual grid, each grid cell capable of holding a single organism. Once the grid is full, new organisms will overwrite random existing ones from the population, creating an evolutionary pressure for rapid replication. To increase replication rate, Avidians compete for increased access to CPU power by consuming virtual resources available in the environment. To consume a resource, organisms must perform an associated logical computation, or ‘task’. Each task can be coupled with a specific resource, enabling adaptive radiation. The amount of CPU power rewarded to a task performing Avidian correlate with the amount of resource consumed, thus rewards decrease as resource abundance is drawn down.

3.2 Environmental configurations and the evolution of digital communities

In this work, the virtual grid was set to 100×100 cells, all of them equally connected, thus creating a well-mixed environment. Organisms were of a fixed length of 100 programming instructions, and point mutation was the only possible mutation type. Each experiment was seeded with a single ancestor organism, capable of self-replication but no tasks, and digital populations evolved at a mutation rate of 0.0025 mutations per copied instruction for 200 000 updates². Because evolution is a stochastic process, communities evolving from identical conditions may develop differently. Therefore, experiments were replicated by evolving 20 different communities under identical conditions. Replicate size was validated by repeating a subset of experiments and checking for equivalent results. Each replicate community was seeded with a unique random seed - a number used to initialize the random number generator that determines the course of evolution in Avida. By specifying the random seed, I make sure experiments are reproducible.

¹<https://avida.devosoft.org/>

²Avida time is measured in updates, each update corresponding to an average of 30 instructions executed per organism. The actual number of instructions executed by a single organism depends on its relative fitness. Organisms with a higher fitness executes faster and will thus execute more than 30 instructions per update.

Experimental environments were initialized with a single resource available. When consumed, this resource created a byproduct also available for consumption, which in turn created yet another byproduct upon consumption, and so on. In total, environments had a potential for 50 resources, 49 of which were created exclusively as byproducts of one another. Conversion rate between resource and byproduct was 1:1, which is equivalent to an environment without any energy loss between trophic levels. This enabled me to study the effects of recurring perturbations isolated from the well-known constraint that energetics impose on food chains (see for example Hutchinson, 1959). Mimicking a chemostat, the initial resource had a continuous inflow of 100 units per update, and all resources had a relative outflow of 10 % per update.

As far as data were available, tasks required for resource consumption were organized by increasing complexity. Organization of the first 25 tasks was based on the frequency at which they evolved in an environment where all resources were available from start. As data were missing for the remaining tasks, they were arbitrarily organized. Organization by increasing complexity was motivated by previous experiments showing that complex tasks are not likely to evolve at all unless simpler tasks are rewarded in the process (Lenski et al., 2003). I validated this assumption by evolving communities in environments where tasks required for resource consumption were randomly ordered (see Appendix A).

3.3 Perturbation experiments

I exposed digital communities that had evolved in unperturbed and well-mixed environments for 200 000 updates, as previously described, to a period of recurring perturbations. Each perturbation event consisted of an instantaneous removal of a predetermined fraction of all organisms in the population. Organisms to be removed were chosen at random. Intensity of perturbations was varied by altering the fraction of organisms removed and the average waiting time between perturbation events. Perturbation intensity (D') was calculated as:

$$D' = 1 - (1 - D)^{\frac{1}{T}}, \quad (1)$$

where D was fraction removed and T was average waiting time between perturbations. Each community was perturbed 100 times, and the timing of perturbations were random, following a Poisson point process. After the perturbation period, communities evolved for an additional 10 000 updates.

In addition to varying perturbation magnitude and rate, I changed the number of perturbations to 50, 100, 200, and 400 events. During these experiments, perturbation intensity was kept constant, with removal of 50 % of all organisms at each perturbation event and an average waiting time of 50 updates between perturbations.

3.4 Phenotype and cross-feeding chain length

Organisms could perform more than one task, thus making it possible for a single organism to participate in cross-feeding interactions on multiple levels of the byproduct chain. For each organism, its phenotype was defined as the number of times it performed each task during one execution of its genetic program. Cross-feeding chain length was then defined as the bottom-most resource it *consumed*, thus requiring both that the organism performed the necessary task and that the resource was available in the environment for it to be considered a part of the organism's cross-feeding chain. The bottom-most resource refers to the resource on the lowest trophic level, starting at the top with the initial resource and moving down the levels with each consecutive byproduct. Thus, if an organism consumed the second and seventh resource, its chain length would be seven.

3.5 Lineage tracing

Of perturbation experiments with at least one surviving community, I saved lineage data of surviving organisms for the four experiments with the highest perturbation intensity. I created phylogenies for each surviving clade, starting from the last common ancestor born before the perturbation period (hereafter referred to as the last common ancestor) and ending with the last descending species³ alive 8 000 updates after the perturbation period stopped (hereafter referred to as tip species, as they constitute the tips, or leaves, of the phylogenetic tree). For each community, I checked which was the lowest level of the byproduct chain maintained through the entire perturbation period, meaning that the resource was continuously supplied as a byproduct from organisms consuming the resource on the level above. Last common ancestors were then defined as long-chained if they consumed resources above that level, and short-chained if they did not. Change in cross-feeding chain length compared to the last common ancestor was calculated for each tip species. During these analyses, all tasks an organism performed were considered part of its cross-feeding chain, regardless of whether the corresponding resource was available in the environment or not.

4 Results

After 200 000 updates, communities in my experiments ($n = 300$) had evolved on average for 6074 ± 1396 generations and consisted on average of 948 ± 390 different phenotypes.

Maximum length of cross-feeding chains in communities evolved in well-mixed and unperturbed conditions ($n = 20$) were 19-35 at the end of the experiment (Figure 1).

³Here, I use species as equivalent to genotype. That is, as soon as a mutation is introduced in the genome of an offspring, I consider that offspring a separate species.

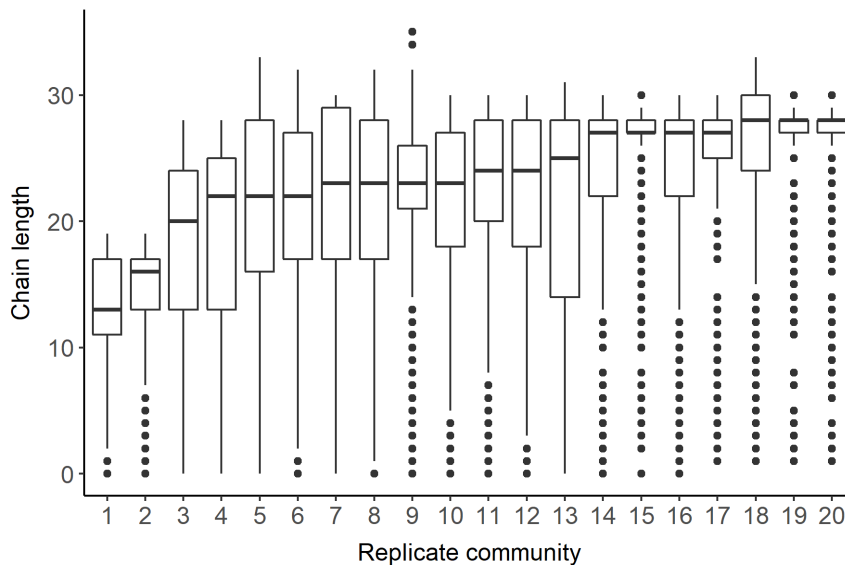


Figure 1: Boxplot showing cross-feeding chain lengths of digital organisms in communities that had evolved for 200 000 updates in well-mixed and unperturbed environments. Maximum 50 resources were available for consumption, 49 of which were formed exclusively as byproducts of another resource. Chain length was defined for each organism as the bottom-most resource it consumed.

During experiment time, one community briefly utilized 47 of the 50 available resources, but that chain was not maintained for any prolonged period of time.

In some digital communities, the recurring perturbations caused all organisms in the population to die, thus resulting in community extinction. Such extinctions increased in frequency with increasing perturbation intensity (Figure 2A). In two experiments, with perturbation intensity 0.0317 and 0.0228 respectively, all replicates resulted in community extinction. In experiments where perturbation intensity was kept constant at 0.0138 and the number of perturbations were varied, community extinctions were rare when perturbation count was 50 (5 % of replicate communities; Figure 3A). When perturbation count was 100, 200, and 400, community extinction occurred in 35, 55, and 50 % of replicates, respectively.

In surviving communities, recurring perturbations caused shortened cross-feeding chains with increased frequency as perturbation intensity increased (i.e., as extinction rate increased or as average waiting time between perturbations decreased; Figure 2B). Perturbation intensity of 0.0152, 0.0172 and 0.0182 caused a distinct shortening of cross-feeding chains in 100, 80, and 100 % of surviving communities, respectively. However, when perturbation intensity was 0.0182, community extinctions were frequent (90 % of communities) and thus results are based on only two surviving communities. When perturbation intensity was low (0.0021 and 0.0071), there was no effect on cross-feeding chains. In some communities, cross-feeding chains remained shortened after perturbations stopped.

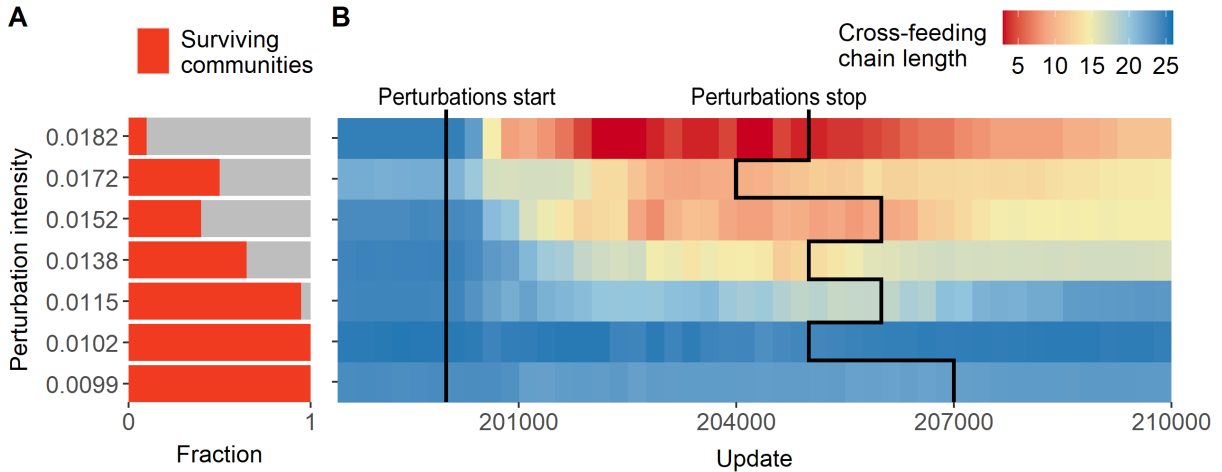


Figure 2: **A** shows the fraction of digital communities ($n = 20$) that survived a period of recurring perturbations, during which a proportion of all organisms were killed at 100 random points in time. Y-axis shows the perturbation intensity; a joined measure of perturbation rate and magnitude (Eq. 1). **B** shows length of cross-feeding chains in each corresponding experiment. Experiment time is shown on x-axis, and color in each tile represents the mean of median chain length in surviving communities at that time (by steps of 250 updates). Chain lengths were defined for each organism as the bottom-most resource it consumed. Start and end of perturbation periods are marked with black lines.

Such persisting effects on chain length were more frequently observed as perturbation intensity increased.

When keeping the perturbation intensity constant at 0.0138, the frequency of shortened cross-feeding chains increased with increasing number of perturbations (Figure 3B). However, even few (50) extinction events caused shortened chains in 32 % of surviving communities. In two of those communities, chain lengths had not recovered to original values by the end of the experiment. As perturbation count increased, such persisting effects on the length of cross-feeding chains became more frequent among surviving communities.

In the four experiments where lineages of surviving organisms were tracked, organisms descending from a long-chained ancestor more frequently evolved shorter chains compared to their ancestor than did organisms descending from a short-chained ancestor. One example is presented in Figure 4, showing two surviving clades from a community exposed to perturbations with an intensity of 0.0172. In the clade descending from a long-chained ancestor, all tip species ('leaves' in the phylogeny) had evolved a shorter chain than their last common ancestor ('root' in the phylogeny), which was born before the perturbation period started. By contrast, a majority of organisms descending from a short-chained ancestor had evolved longer chains than their last common ancestor. More examples are available in Appendix B.

Mean change in cross-feeding chain length within a clade negatively correlated with chain length of the last common ancestor of that clade (Figure 5). The pattern was

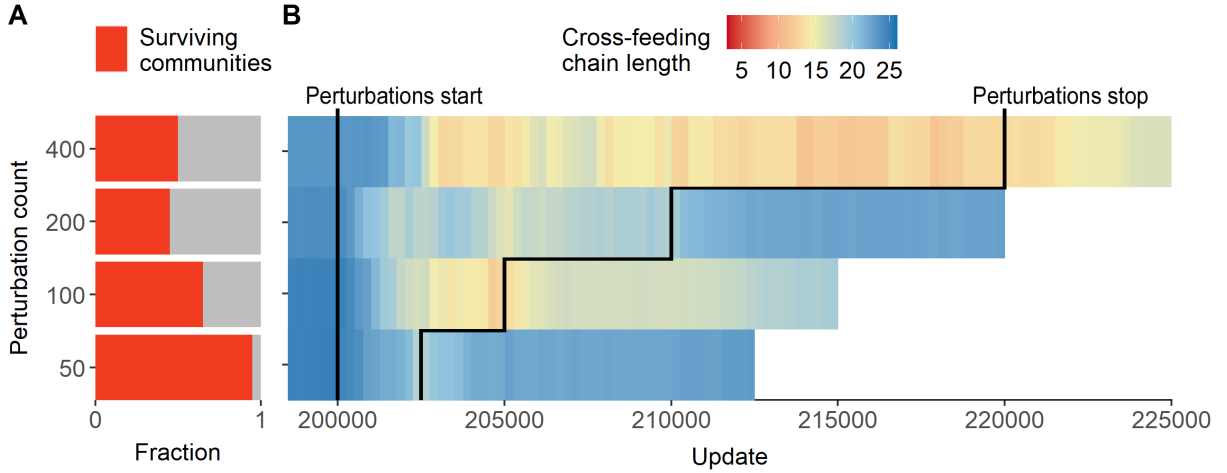


Figure 3: **A** shows the fraction of digital communities ($n = 20$) that survived a period of recurring perturbations, during which a proportion of all organisms were killed at random points in time. Perturbation intensity (Eq. 1) was 0.0138, and perturbation count is shown on y-axis. **B** shows length of cross-feeding chains in each corresponding experiment. Experiment time is shown on x-axis, and color in each tile represents the mean of median chain length in surviving communities at that time (by steps of 250 updates). Chain length was defined for each organism as the bottom-most resource it consumed. Start and end of perturbation periods are marked with black lines.

consistent independent of perturbation intensity. However, only two communities survived perturbations with an intensity of 0.0182, each of them consisting of a single clade, thus resulting in very few data points. It is also important to note here that these data consist of clades from all surviving communities, including ones not showing any decrease in cross-feeding chain length during the perturbation period.

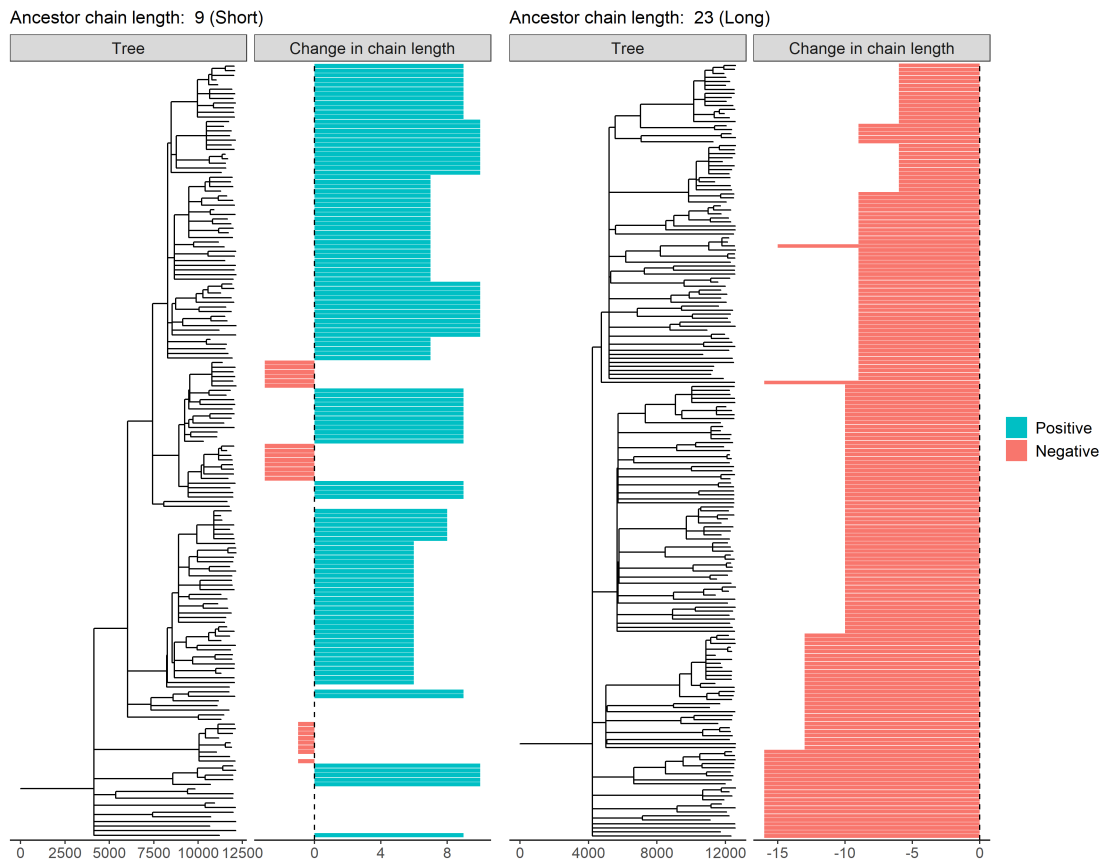


Figure 4: Example of two clades descending from a short-chained and a long-chained ancestor respectively. Change in cross-feeding chain length compared to the last common ancestor (the 'root') within each clade is shown as a bar for each genotype constituting a tip ('leaf') in the phylogeny. Blue bars indicate an increase in cross-feeding chain length and red bars indicate a decrease.

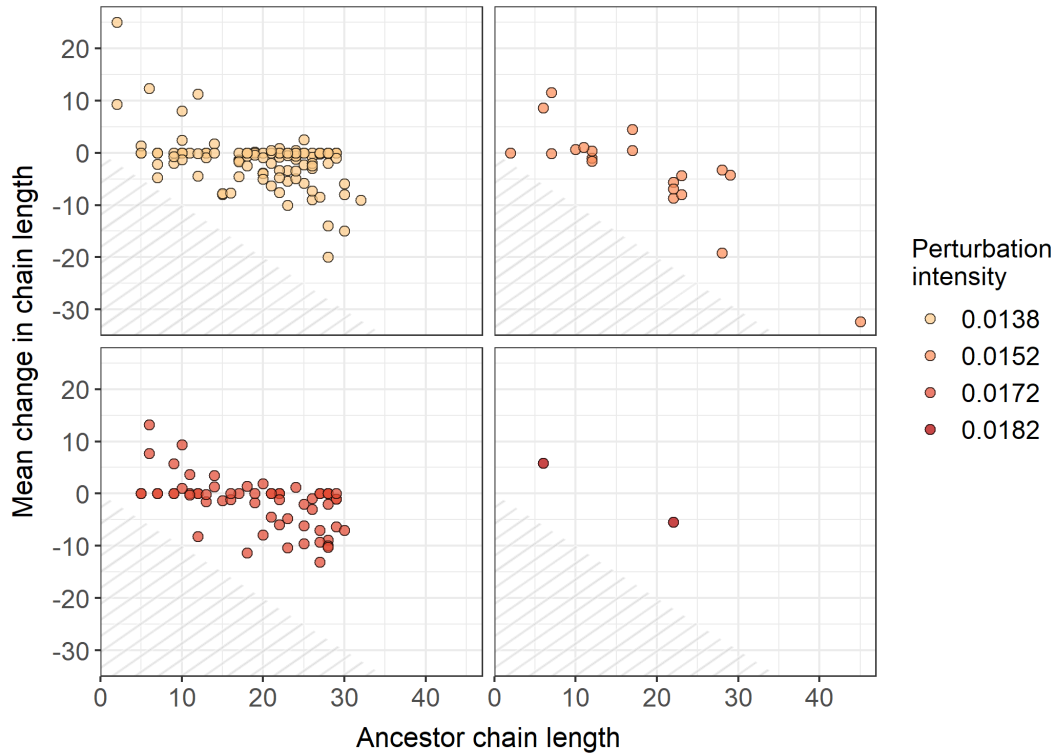


Figure 5: Change in cross-feeding chain length when comparing species alive 8 000 updates after a period of recurring perturbations with their last common ancestor, born before perturbations started. Each point represents the mean change of all species within a clade which constitute a tip in the phylogenetic tree. Color indicates the intensity of perturbations which the communities were exposed to, each intensity shown in a separate facet. Striped areas indicate restrictions to values of change in chain length, as trait value was strictly non-negative.

5 Discussion

In this study, populations of digital organisms regularly evolved cross-feeding chains of 30 steps or more in environments unconstrained by energy loss between trophic levels, showing that there are no inherent constraints to evolving long chains of byproduct cross-feeding. Exposing cross-feeding communities to a period of recurring perturbations, during which a fraction of all organisms was removed from the population at random points in time, reduced the length of cross-feeding chains, especially if perturbations were strong or frequent. In some communities, effects on cross-feeding chains persisted after the perturbation period had ended, possibly due to evolutionary adaptations causing a species-level loss of abilities to consume some resources.

Understanding the human gut microbiome assembly and functioning is of great medical importance, and many studies have contributed to our knowledge of associated microbes and metabolites (see for example Turnbaugh et al., 2007; Shafquat et al., 2014; Sung et al., 2017). In a recent study, Wang and colleagues (2019) used a model framework

to estimate the metabolic flow of the gut microbiome. They found that the numerous cross-feeding interactions can be represented as movement of metabolites between on average four trophic levels. However, experiment-based mappings of metabolic capabilities of microbiome-associated species reveal that more levels should be possible. In addition to previously known factors limiting the number of trophic levels, such as energetics (Hutchinson, 1959) and population dynamics (Pimm and Lawton, 1977), the movement and limited length of the gut has been suggested as a constraint on length of food chains in the gut microbiome (Wang et al., 2019). The result of this study indicates that recurring perturbations, for example caused by bowel movement, has potential to add yet another constraint on the number of trophic levels in the gut microbiome. Stronger perturbations, caused by for example diarrhea (Pop et al., 2014) or antibiotic treatment (Dethlefsen et al., 2008), which is known to cause long-lasting changes in community composition (Dethlefsen and Relman, 2011), could possibly have even larger effects on the microbiome cross-feeding network.

Though the length of cross-feeding chains in this study consistently decreased when communities were exposed to perturbations of high intensity, there were little or no effect on communities when perturbation intensity was low. However, the number of perturbations in this study was relatively low, and it is possible that a longer period of perturbations would have caused an effect even at low perturbation intensity. A benefit of a digital study system is the ability to test this directly, by exact replication of experiments up to a given time, at which point you expose them to perturbations of different strength or number. In this study, when perturbation intensity was kept constant and perturbation count was varied between 50-400, increasing the number of perturbations increased the frequency of shortened cross-feeding chains, but did not increase the number of community extinctions. While this might suggest that communities adapt to the recurring perturbations, causing community extinctions to become less frequent towards the middle and end of the perturbation period, it might also be a result of the stochastic nature of eco-evolutionary experiments. To draw any general conclusions on the effects of the number of perturbations, further experiments are needed across multiple perturbation intensities.

Adaptations can arise quickly in microbial communities (Good et al., 2017), and it is thus relevant to study effects on the gut microbiome from an eco-evolutionary perspective. Using the Avida software for digital evolution granted access to a complete record of the course of evolution. When tracking lineages of species surviving the period of recurring perturbations, results indicated that organisms descending from ancestors with long cross-feeding chains more often lost the ability to consume resources late in the byproduct chain than did organisms descending from short-chained ancestors. On the opposite, organisms descending from short-chained ancestors commonly evolved new abilities to consume resources late in the byproduct chain. This suggests that exposing cross-feeding

communities to recurring perturbations can cause an evolutionary pressure to consume resources early in the byproduct chain. A possible explanation for this is that, at each level of the chain, abundance of the byproduct resource depends on consumption of the resource on the level above, creating a chain of interdependent cross-feeding interactions. If the chain would be interrupted, the byproduct resource below that point would soon be depleted. This suggests that the further down the byproduct chain an organism consumes resources, the higher is the risk of its resource being depleted.

The species-level loss of abilities to consume resources may explain the delayed recovery of cross-feeding chain length on a community level, as tasks required for resource consumption must evolve anew. Similar effects were found during a study where digital communities were exposed to a press perturbation, consisting of a period with low abundance of the primary resource (Yedid et al., 2009). The press perturbation caused a selection for short generation times and a loss of trophic functions, which in turn caused a delayed recovery to the original resource consumption, compared to communities exposed to a pulse perturbation. To conclude if similar mechanisms are causing the persisting effects on cross-feeding chains observed in this study, further quantitative analyses are needed to investigate if the observed pattern is consistent over perturbation intensities. Further experiments are also necessary to determine the effects of shortened cross-feeding chains on community function and resilience. For example, would communities exposed to a period of recurring perturbations be less sensitive to a second period of perturbations than a community evolved in strictly unperturbed conditions?

Using a digital study system enabled a separation of different factors that would not have been feasible in a biological system, most importantly by allowing a conversion between resource and byproduct without any loss of energy. In future studies, reintroducing energy loss between trophic levels will enable investigations of the combined effects of constraints imposed by energetics and recurring perturbations. It is possible that communities evolving in such an environment will have a different composition than communities in this study, and thus might respond differently to perturbations.

5.1 Conclusions

Recurring perturbations have potential to constrain the length of cross-feeding chains in environments with complete energy transfer between trophic levels. In communities of digital organisms, cross-feeding chains commonly reached 30 steps or more, which is one order of magnitude longer than the average chain length in the human gut. When exposing these communities to recurring perturbations, cross-feeding chains regularly shortened, especially when perturbations were strong or frequent. Some evidence suggests that organisms adapted to recurring perturbations by losing the ability to consume resources late in the byproduct chain, which could explain the persisting effects on chain length observed in some communities after perturbations stopped. While it is evident that

perturbations constrain cross-feeding chains in this environment, further experiments, for example on effects of energy loss between trophic levels, are needed to expand conclusions to biological systems, such as the gut microbiome. It is also necessary to investigate under which circumstances perturbations have an effect and not.

6 Societal and ethical considerations

The human gut microbiome is of immense importance for normal health and development (Brestoff and Artis, 2013; Cho and Blaser, 2012). Changes in community composition have been linked to multiple diseases, such as diabetes (Qin et al., 2012) and inflammatory bowel disease (Franzosa et al., 2019). Manipulation of the gut microbiome has great potential in both preventing and treating such diseases, but to make use of this potential, we must increase our understanding of both the ecological and evolutionary mechanisms shaping the microbiome composition and functioning. A recent study have shown that the numerous chains of cross-feeding interactions in the human gut are on average four steps long (Wang et al., 2019), despite the presence of species with metabolic capabilities to continue these chains (Sung et al., 2017). What constrains these chains are not yet known, and in this study I investigated if recurring perturbations can limit the length of such cross-feeding chains. This study contributes to our overall understanding of the human gut microbiome functioning, and thus to its role in human health. Though the scope of the study is to look specifically at recurring perturbations, its results have implications for studying other possible perturbations, such as diarrhea or antibiotic treatment.

Using a digital study system has many benefits compared to a biological one, some of them mentioned in the introduction and discussion. Other benefits include reduced risks compared to working with actual microbes, many of them pathogens, and reduced resource use, as experiments are much faster than when using biological systems. Expanding the use of digital evolution in ecology research has a lot of potential in developing new and effective frameworks for studying for example population dynamics or evolutionary responses to abiotic changes.

7 Acknowledgments

I would like to thank my supervisor György Barabás for his guidance through this project. I am also grateful to Akshit Goyal for inspiring the topic of this project, as well as for valuable discussions about the results. I thank Luis Zaman for sharing his great knowledge of Avida, and for his help with interpreting results and providing data on evolution of computational tasks. I would also like to thank my examiner Anna Eklöf and reviewer Rebecca Petersen for their helpful comments when writing this thesis.

References

- Brestoff, J. R., Artis, D., 2013. Commensal bacteria at the interface of host metabolism and the immune system. *Nature Immunology* 14, 676–684.
- Cho, I., Blaser, M. J., 2012. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* 13, 260–270.
- Chow, S. S., Wilke, C. O., Ofria, C., Lenski, R. E., Adami, C., 2004. Adaptive radiation from resource competition in digital organisms. *Science* 305, 84–86.
- Cooper, T. F., Ofria, C., 2002a. Evolution of stable ecosystems in populations of digital organisms. In: *Artificial Life VIII: Proceedings of the Eighth International Conference on Artificial life*. pp. 227–232.
- Cooper, T. F., Ofria, C., 2002b. Evolution of stable ecosystems in populations of digital organisms. In: *Artificial Life VIII: Proceedings of the Eighth International Conference on Artificial life*. pp. 227–232.
- Dethlefsen, L., Huse, S., Sogin, M. L., Relman, D. A., 11 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16s rRNA sequencing. *PLOS Biology* 6 (11), 1–18.
- Dethlefsen, L., Relman, D. A., 2011. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences* 108, 4554–4561.
- Elena, S. F., Wilke, C. O., Ofria, C., Lenski, R. E., 2007. Effects of population size and mutation rate on the evolution of mutational robustness. *Evolution* 61, 666–674.
- Franzosa, E. A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H. J., Reinker, S., Vatanen, T., Hall, A. B., Mallick, H., McIver, L. J., Sauk, J. S., Wilson, R. G., Stevens, B. W., Scott, J. M., Pierce, K., Deik, A. A., Bullock, K., Imhann, F., Porter, J. A., Zhernakova, A., Fu, J., Weersma, R. K., Wijmenga, C., Clish, C. B., Vlamakis, H., Huttenhower, C., Xavier, R. J., 2019. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nature Microbiology* 4, 293–305.
- Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., Desai, M. M., 2017. The dynamics of molecular evolution over 60,000 generations. *Nature* 551 (7678), 45–50.
- Hutchinson, G. E., 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? *American Naturalist* 93 (870), 137–145.

- Lehman, J., Clune, J., Misevic, D., Adami, C., Altenberg, L., Beaulieu, J., Bentley, P. J., Bernard, S., Beslon, G., Bryson, D. M., Chrabaszcz, P., Cheney, N., Cully, A., Doncieux, S., Dyer, F. C., Ellefsen, K. O., Feldt, R., Fischer, S., Forrest, S., Frénoy, A., Gagné, C., Goff, L. L., Grabowski, L. M., Hodjat, B., Hutter, F., Keller, L., Knibbe, C., Krcak, P., Lenski, R. E., Lipson, H., MacCurdy, R., Maestre, C., Miikkulainen, R., Mitri, S., Moriarty, D. E., Mouret, J.-B., Nguyen, A., Ofria, C., Parizeau, M., Parsons, D., Pennock, R. T., Punch, W. F., Ray, T. S., Schoenauer, M., Shulte, E., Sims, K., Stanley, K. O., Taddei, F., Tarapore, D., Thibault, S., Weimer, W., Watson, R., Yosinski, J., 2018. The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities. In: *Artificial Life Conference Proceedings*. MIT Press, pp. 55–56.
- Lenski, R. E., Ofria, C., Collier, T. C., Adami, C., 1999. Genome complexity, robustness and genetic interactions in digital organisms. *Nature* 400, 661–664.
- Lenski, R. E., Ofria, C., Pennock, R. T., Adami, C., 2003. The evolutionary origin of complex features. *Nature* 423 (6936), 139–144.
- Lewontin, R. C., 1970. The units of selection. *Annual Review of Ecology and Systematics* 1, 1–18.
- Ofria, C., Wilke, C. O., 2004. Avida: A software platform for research in computational evolutionary biology. *Artificial Life Models in Software (Second Edition)* 229, 3–35.
- Pennock, R. T., 2007. Models, simulations, instantiations, and evidence: The case of digital evolution. *Journal of Experimental and Theoretical Artificial Intelligence* 19 (1), 29–42.
- Pimm, S. L., Lawton, J. H., 1977. Number of trophic levels in ecological communities. *Nature* 268, 329–331.
- Pop, M., Walker, A. W., Paulson, J., Lindsay, B., Antonio, M., Hossain, M. A., Oundo, J., Tamboura, B., Mai, V., Astrovskaya, I., Bravo, H. C., Rance, R., Stares, M., Levine, M. M., Panchalingam, S., Kotloff, K., Ikumapayi, U. N., Ebruke, C., Adeyemi, M., Ahmed, D., Ahmed, F., Alam, M. T., Amin, R., Siddiqui, S., Ochieng, J. B., Ouma, E., Juma, J., Mailu, E., Omore, R., Morris, J. G., Breiman, R. F., Saha, D., Parkhill, J., Nataro, J. P., Stine, O. C., 2014. Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biology* 15, R76.
- Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., Peng, Y., Zhang, D., Jie, Z., Wu, W., Qin, Y., Xue, W., Li, J., Han, L., Lu, D., Wu, P., Dai, Y., Sun, X., Li, Z., Tang, A., Zhong, S., Li, X., Chen, W., Xu, R., Wang,

- M., Feng, Q., Gong, M., Yu, J., Zhang, Y., Zhang, M., Hansen, T., Sanchez, G., Raes, J., Falony, G., Okuda, S., Almeida, M., LeChatelier, E., Renault, P., Pons, N., Batto, J.-M., Zhang, Z., Chen, H., Yang, R., Zheng, W., Li, S., Yang, H., Wang, J., Ehrlich, S. D., Nielsen, R., Pedersen, O., Kristiansen, K., Wang, J., 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60.
- Shafquat, A., Joice, R., Simmons, S. L., Huttenhower, C., 2014. Functional and phylogenetic assembly of microbial communities in the human microbiome. *Trends in Microbiology* 22 (5), 261–266.
- Sung, J., Kim, S., Cabatbat, J. J. T., Jang, S., Jin, Y., Jung, G. Y., Chia, N., Kim, P., 2017. Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nature Communications* 8, 15393.
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., Gordon, J. I., 2007. The human microbiome project. *Nature* 449, 804–810.
- Wang, T., Goyal, A., Dubinkina, V., Maslov, S., 2019. Evidence for a multi-level trophic organization of the human gut microbiome. *PLOS Computational Biology* 15, 1–20.
- Yedid, G., Ofria, C. A., Lenski, R. E., 2009. Selective press extinctions, but not random pulse extinctions, cause delayed ecological recovery in communities of digital organisms. *The American Naturalist* 173 (4), E139–E154.

Appendix

A Task order

To test the assumption that tasks should be ordered by increasing complexity to enable evolution of byproduct cross-feeding chains, I evolved 20 communities of digital organisms in an environment where tasks required for resource consumption were randomly ordered. Task order was different for each community. In these environments, communities evolved much shorter chains than in environments where tasks were ordered by increasing complexity. This is in agreement with previous studies, which have found that complex tasks are more likely to evolve if simple tasks are rewarded in the process (Lenski et al., 2003).

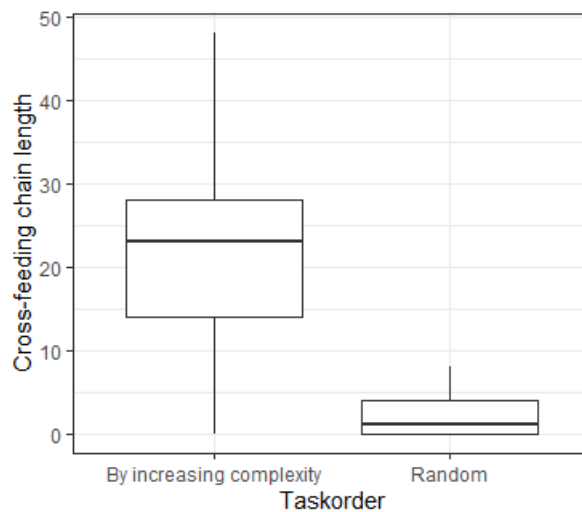


Figure A.1: Distribution of cross-feeding chain lengths in experiments with tasks ordered by increasing complexity vs. random task order. Each box represents 20 replicates.

In many experiment communities, cross-feeding chains of exactly 30 levels evolved. However, a few communities evolved longer chains, some of them utilizing all 50 resources, showing that there is no inherent constraint to evolving such long chains. I hypothesized that the task required for consuming the resource on level 31 was especially complex, thus restricting evolution of longer chains in many communities. To test this, I evolved 20 communities in environments where the task usually required for resource consumption on level 31 was instead required for consuming the resource on level 20. Of communities evolved in this environment, 70 % evolved chains of length 19. In environments with the original task order, only 10 % of communities had chains of length 19, and all other communities evolved chains of at least 28 levels. Thus, the result support the hypothesis that the reason chains so often stopped at level 30 was because the task required for consuming the next byproduct resource was especially complex and thus less likely to evolve.

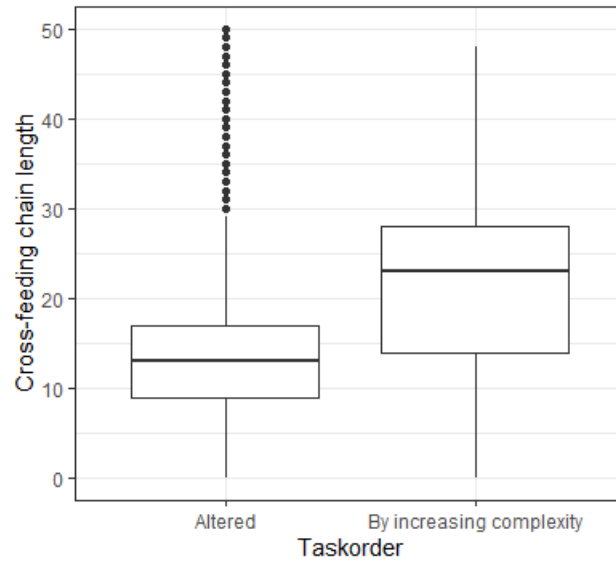


Figure A.2: Distribution of cross-feeding chain lengths in experiments with tasks in the original order vs. when the task required for resource consumption on level 31 had been interchanged with the task required on level 20 (altered). Each box represents 20 replicates.

B Evolution of chain length in perturbed communities

Lineages of all organisms were traced in communities perturbed with intensities 0.0138, 0.0152, 0.0172, and 0.0182. Last common ancestors born before perturbations started were for each clade defined as having either a long or short cross-feeding chain (for further detail, see Methods). Many surviving communities had only a single surviving clade, or contained clades descending from only short-chained or only long-chained organisms. However, six communities contained at least one clade each descending from a long-chained and short-chained ancestor, here shown with phylogenies and data of chain change between ancestors and tip species for each clade.

In general, organisms descending from long-chained ancestors were more frequently observed to evolve shorter chains than longer chains. On the opposite, organisms descending from short-chained ancestors often evolved longer cross-feeding chains (see Figures B.1, B.3, B.4, and B.5). Though some exceptions occurred (see Figures B.2 and B.6), the pattern was consistent among most communities (for an overview, see Results, for example Figure 5).

Perturbation intensity: 0.0152

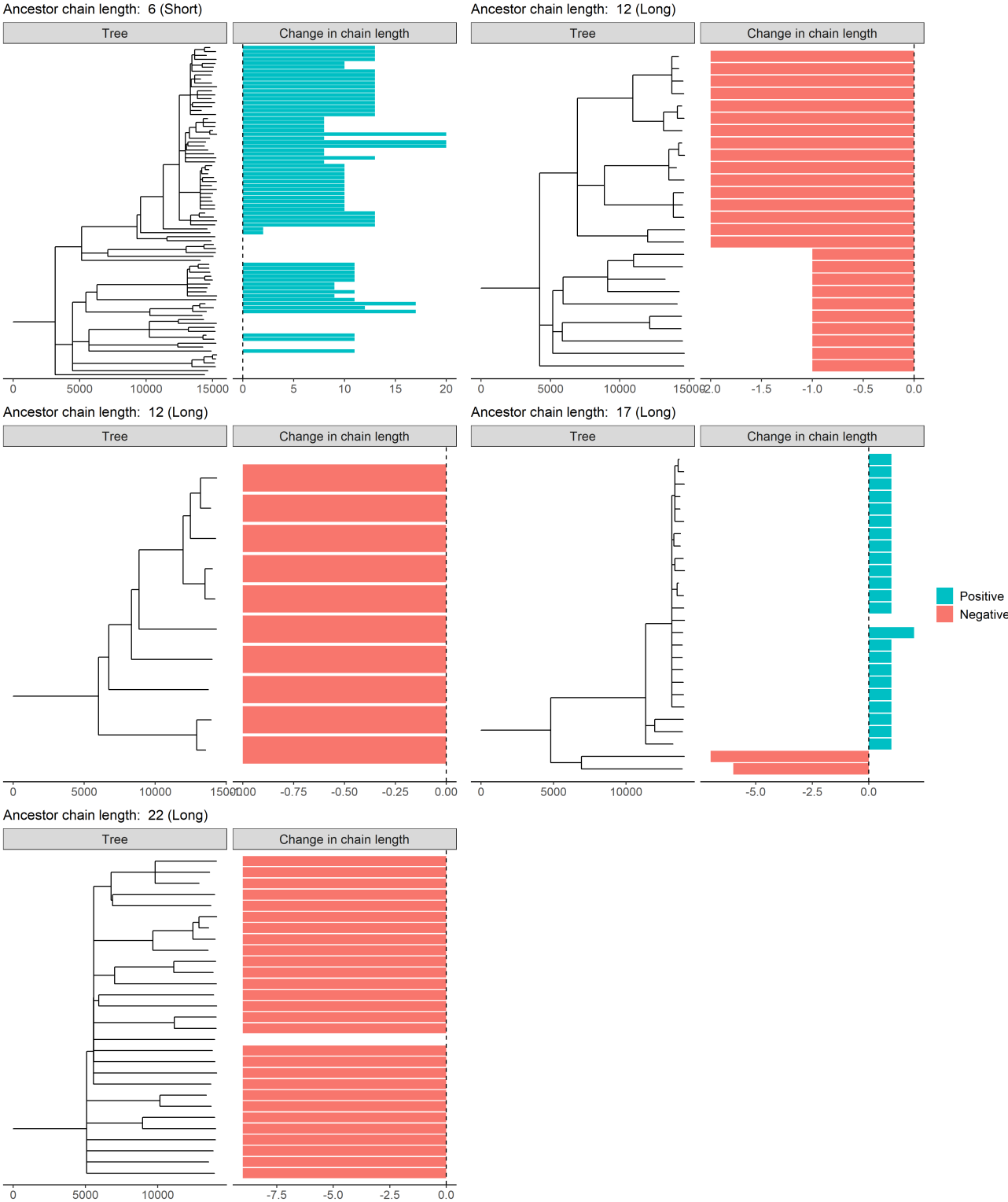


Figure B.1: Phylogenies and change in chain length in community 1 exposed to perturbation of intensity 0.0152.

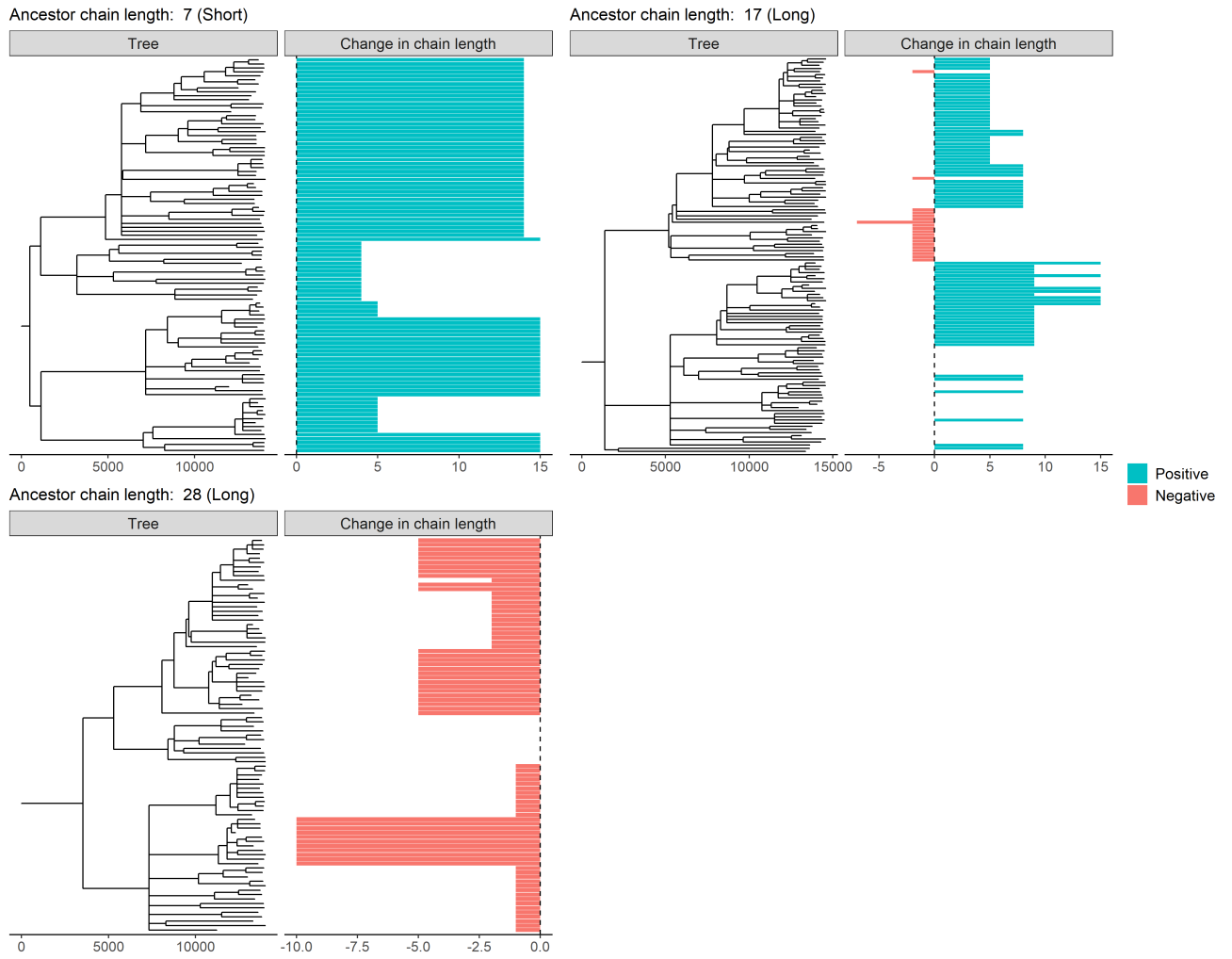


Figure B.2: Phylogenies and change in chain length in community 2 exposed to perturbation of intensity 0.0152.

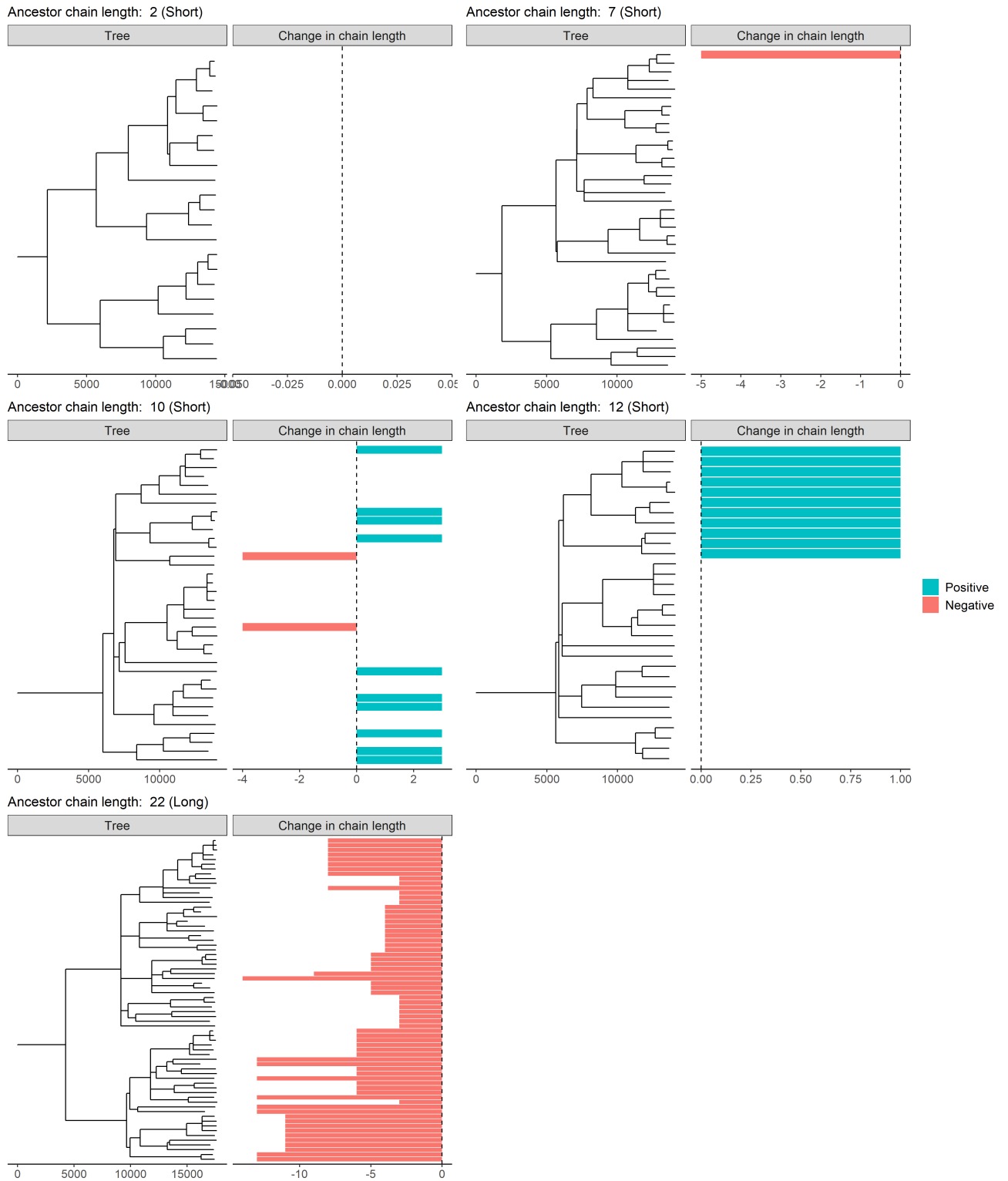


Figure B.3: Phylogenies and change in chain length in community 3 exposed to perturbation of intensity 0.0152.

Perturbation intensity: 0.0172

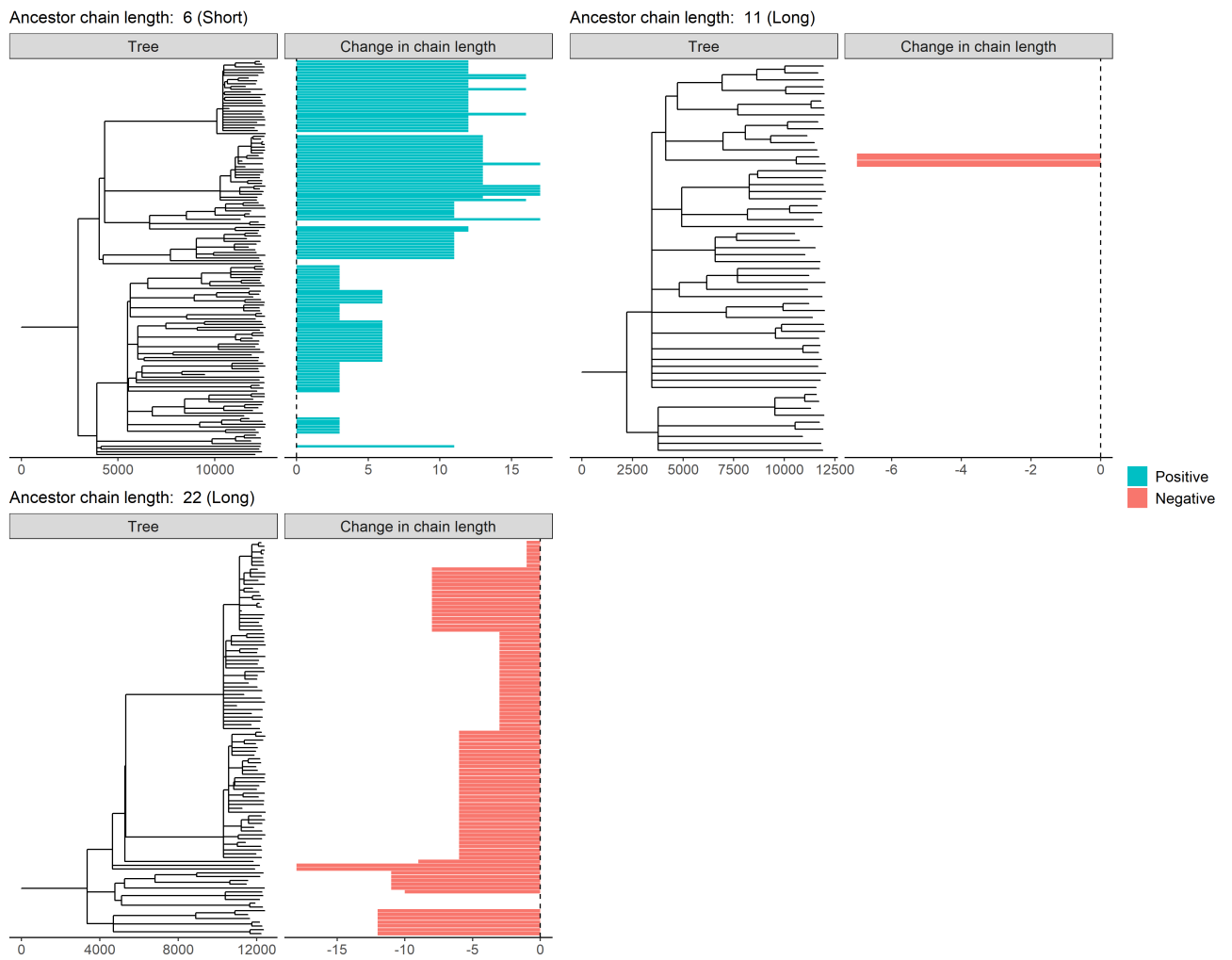


Figure B.4: Phylogenies and change in chain length in community 1 exposed to perturbation of intensity 0.0172.

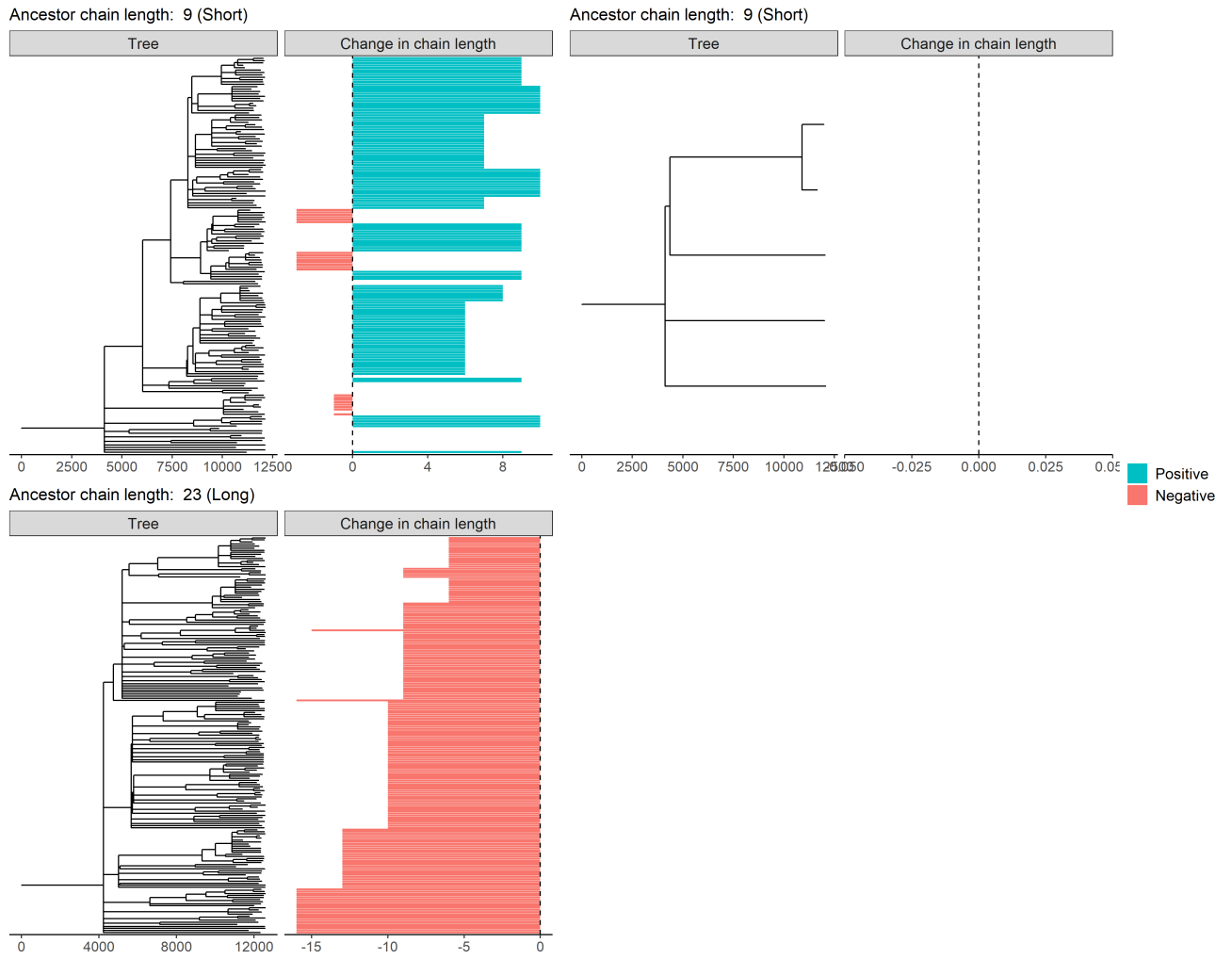


Figure B.5: Phylogenies and change in chain length in community 2 exposed to perturbation of intensity 0.0172.

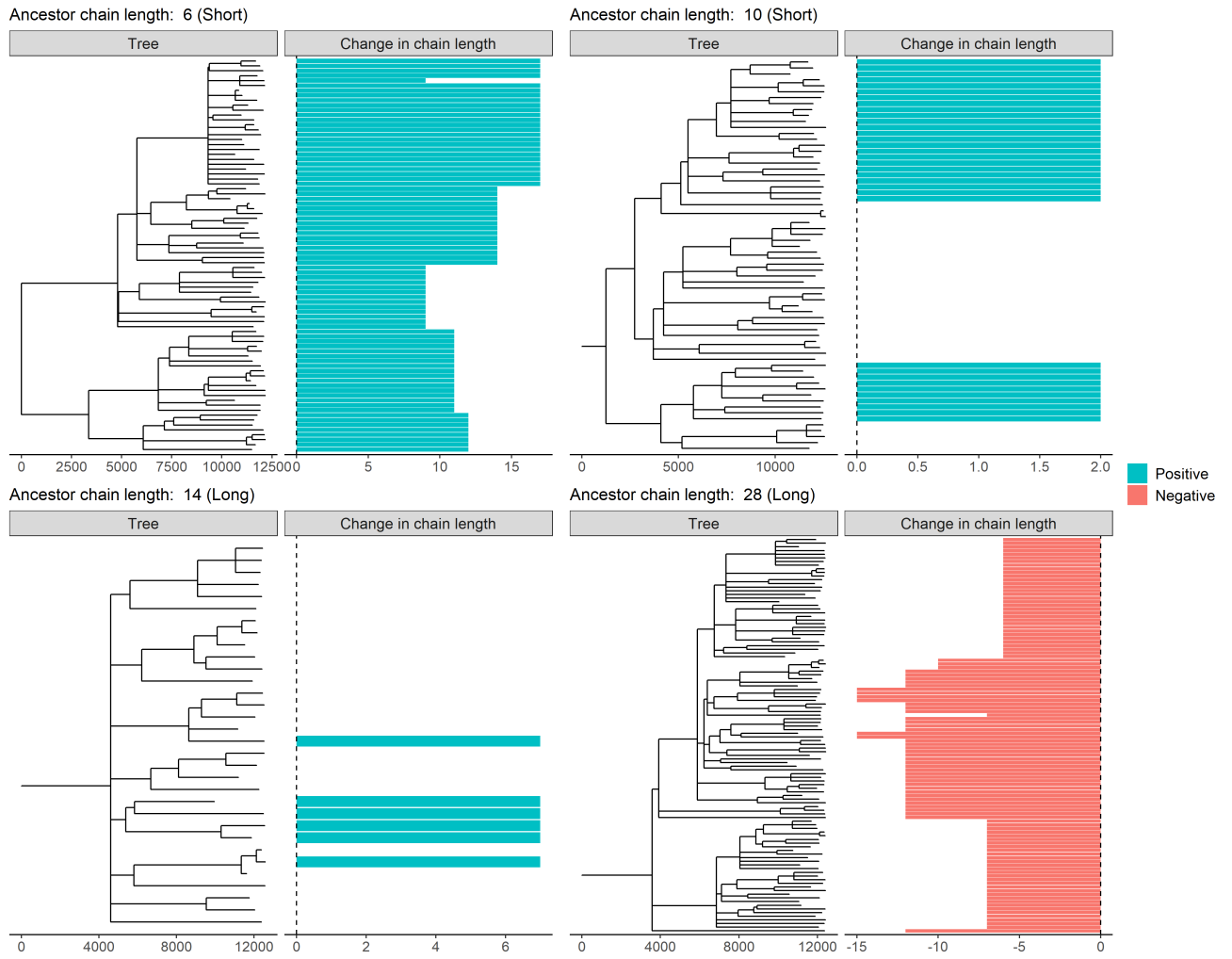


Figure B.6: Phylogenies and change in chain length in community 3 exposed to perturbation of intensity 0.0172.