Eco-logical autocalysis: a central principle in ecosystem organization?

MICHEL P. VELDHUIS,1,4 MATTY P. BERG,1,2 MICHEL LOREAU,3 AND HAN OLFF1

1 Groningen Institute for Evolutionary Life Sciences, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands
2 Department of Ecological Science, Vrije Universiteit, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
3 Centre for Biodiversity Theory and Modeling, Theoretical and Experimental Ecology Station, CNRS and Paul Sabatier University, 09200 Moulis, France

Abstract. Ecosystems comprise flows of energy and materials, structured by organisms and their interactions. Important generalizations have emerged in recent decades about conversions by organisms of energy (metabolic theory of ecology) and materials (ecological stoichiometry). However, these new insights leave a key question about ecosystems inadequately addressed: are there basic organizational principles that explain how the interaction structure among species in ecosystems arises? Here we integrate recent contributions to the understanding of how ecosystem organization emerges through ecological autocalysis (EA), in which species mutually benefit through self-reinforcing circular interaction structures. We seek to generalize the concept of EA by integrating principles from community and ecosystem ecology. We discuss evidence suggesting that ecological autocalysis is facilitated by resource competition and natural selection, both central principles in community ecology. Furthermore, we suggest that pre-emptive resource competition by consumers and plant resource diversity drive the emergence of autocalytic loops at the ecosystem level. Subsequently, we describe how interactions between such autocalytic loops can explain pattern and processes observed at the ecosystem scale, and summarize efforts to model different aspects of the phenomenon. We conclude that EA is a central principle that forms the backbone of the organization in systems ecology, analogous to autocalytic loops in systems chemistry.

Key words: autocalytic loops; community ecology; ecosystem ecology; interaction structure; positive feedback; pre-emptive resource competition; resource diversity; self-organization.

INTRODUCTION

The systems biology approach has radically changed the fields of biochemistry, cell biology, and organismal physiology in recent decades (Hartwell et al. 1999, Kitano 2002, Raes and Bork 2008, Keurentjes et al. 2011). The recognition that non-living dynamic systems can exhibit complex and self-organizing emergent behavior that are not simply predictable from the properties of their basic building blocks (Levin 1998, 1999, Sole and Goodwin 2008) inspired many scientists to (re-)examine the organization of interactions in cells and organisms, rather than to focus on the characteristics of isolated components, such as genes (Rosen 1991, Kitano 2002, Sun and Becksé 2010). Preceding this trend by decades, the field of ecosystem ecology was one of the first biological disciplines to adopt a complex systems perspective by mapping and quantifying fluxes of energy and nutrients among biotic and abiotic compartments (Lindeman 1942, Odum 1953, 1968, Pace and Groffman 1998), and describing them as circular causal systems (Hutchinson 1948). But despite this early start, and the revolutions inspired by a complex (adaptive) systems perspective in cell biology, genetics and developmental biology, the field of ecosystem ecology still has few generally accepted principles for how the organization of species interactions comes about. If general principles of ecosystem organization exist, they are expected to be found as regularities in the structure of the hybrid network of different types of interactions among species: predation, competition, mutualism, parasitism, and ecosystem engineering (Olff et al. 2009). Integrative steps, based on knowledge of the dynamics of evolutionary interactions, are now needed to develop satisfying general concepts of ecosystem structure and organization (Levin et al. 2001, Loreau et al. 2001, Naeem 2002, Hooper et al. 2005, Loreau 2010a).

So far, the food web approach comes closest in examining general rules of ecosystem organization (Ings et al. 2009, Thompson et al. 2012), and provides a valuable starting point. However, most present-day food web approaches lack specific key elements of whole-ecosystem organization. First, food web studies and models generally do not incorporate non-trophic interactions, despite these having been shown to be very important for ecosystem organization and functioning (Goudard and Loreau 2008, Olff et al. 2009, Kéfi et al. 2012, 2015). Second, most food web studies assume unidirectional flows of energy and matter while ecosystems are characterized by cycles, especially for limiting nutrients (Patten and Odum 1981; but see Loreau 2010b). Third, food web studies often describe only parts of ecosystems or modules (e.g., plant–
Ecology Bivalves and their endosymbiotic sulfide-oxidizing gill bacteria profit from seagrasses through native configurations, and may even explain the origin of life often statistically and thermodynamically favored over alternative and especially more open configurations. The conceptual relationship between chemical stoichiometry and the highly successful field of ecological stoichiometry (Elser et al. 2000, Sterner and Elser 2002) is the same as between chemical autocatalysis and ecological autocatalysis. Ecological stoichiometry deals with basic principles in organismal-level conversion of energy and materials, while ecological autocatalysis deals with basic principles in ecosystem-level organization of energy flows and nutrient cycling.

Autocatalysis

Several studies now suggest that the “core engine” of many (if not all) ecosystems is formed by an autocatalytic set of species populations that promote each other in a loop through positive feedbacks (Table 1). From a systems biology perspective, such autocatalytic loops or sets are not at all new. Autocatalytic sets were originally defined in terms of chemical species interacting in biochemical systems, where reactions between interacting species catalyze enough substrate for the next reaction so that the whole set of chemical reactions is self-sustaining given sufficient input of energy and essential materials (Kauffman 1986, Hordijk and Steel 2004, Mossel and Steel 2005). This arises when a set of chemical species form an autocatalytic loop (e.g., A catalyzes the formation of B, B catalyzes C, and C catalyzes the formation of A). Such chemical autocatalytic loops, such as the regular and reverse Krebs cycle, are found at the heart of the intermediate metabolism of all organisms. They are often statistically and thermodynamically favored over alternative configurations, and may even explain the origin of life (Eigen and Schuster 1979, Morowitz et al. 2000, Lincoln and Joyce 2009, Giri and Jain 2012). Hence, they are not trivial or rare structures in biochemical interaction networks. By analogue, interacting populations of different species in an ecosystem can form an autocatalytic loop if each species produces the resources needed by the next species in the loop, in such a configuration that the whole set of species is self-promoting and self-sustaining given sufficient input of energy and essential materials (carbon, nutrients). For example, species A produces the resources needed by species B, B produces the resources needed by C, and C produces the resources needed by A. Similar to biochemical autocatalytic loops in a cell, such ecological autocatalytic loops in ecosystems are expected to be thermodynamically favored over alternative and especially more open configurations. The conceptual relationship between chemical stoichiometry and the highly successful field of ecological stoichiometry (Elser et al. 2000, Sterner and Elser 2002) is the same as between chemical autocatalysis and ecological autocatalysis. Ecological stoichiometry deals with basic principles in organismal-level conversion of energy and materials, while ecological autocatalysis deals with basic principles in ecosystem-level organization of energy flows and nutrient cycling.

A long history in ecosystem ecology

Building on the pioneering work of R. E. Ulanowicz (Ulanowicz 1997, 2009a, Ulanowicz and AbarcaArenas 1997), we explore the idea that sets of biological species that form autocatalytic loops logically arise at the ecosystem level through self-organization. Further, particular configurations of species that promote each other most are most likely to prevail, with community assembly processes (Diamond 1975) and ecosystem assembly processes (Schrama et al. 2012, 2013) forming the underlying dynamics. The work by Ulanowicz builds on initial efforts by Lindeman (1942) and Hutchinson (1948), who were instigators of cybernetic approaches in ecology that explored the origin, structures, constraints, and possibilities of regulatory systems. Subsequently, Odum (1971) advanced their ideas in his descriptions of “reward loops” that led later to recognition of the importance of positive interactions (Bertness and Callaway 1994), positive feedback (DeAngelis et al. 1986), emergent properties (Jørgensen et al. 1992) and formal causality (Ulanowicz 1997). From a theoretical perspective,
Loreau (1998, 2010a) made important contributions using ecological and evolutionary models to investigate the stability and behavior of such circular species configurations and the implications for nutrient cycling. However, despite the early origins and long development of this idea of self-enforcing species interactions, the far-reaching implications of the concept of autocatalytic loops for ecosystem organization are only now gradually emerging (Gadgil and Kulkarni 2009, Ulanowicz 2009a, Hordijk et al. 2012).

The universality of circular interaction structures (loops) in ecology

Nutrient cycling is one of the fundamental aspects of the organization and dynamics of ecosystems. Producers, consumers, and decomposers form interaction structures that comprise circuits (loops) of chemical elements. Such structures form the backbone of ecosystem organization (Hutchinson 1948, Odum 1960, 1968) and key concepts such as ecological stoichiometry (Elser et al. 2000, Sterner and Elser 2002). A simple example helps to illustrate this. Imagine a grass species that grows using mineral nitrogen taken up from the soil. In doing so, it produces detritus that stimulates the growth of a population of a particular detritivorous bacteria that, in turn, makes nitrogen again available for the plants through organic matter decomposition (Fig. 1A), and thus recycles it to the mineral pool. This circular causal relationship between autotrophs that fix energy in complex organic molecules and decomposers that break these down, recycling the mineral nutrients required by autotrophs, forms the basic structure of all ecosystems and is the generic motor on which all life on earth depends. Subsequently, we can extend this loop by adding an earthworm species that consumes the litter. Through fragmentation and mixing of litter, the earthworm stimulates bacterial growth and thus making a pre-processing step before bacterial decomposition (Fig. 1B; Ingham et al. 1985, Standing et al. 2006). Note that this loop of species cannot simply be described by classic predator-prey interactions, which would miss the elements of ecosystem engineering (bioturbation by earthworms, facilitation of microbes by earthworms) and nutrient recycling. Consumers in this case do not negatively affect the growth rate of their resource (as in standard predator–prey interactions), but instead stimulate the production of their food by benefiting other species and thus indirectly benefiting other species and themselves. The result is a circular interaction configuration (loop) based on a set of consumer–resources linkages through which both energy and nutrients flow. These configurations can be further extended by species that insert at different positions in the loop, similar to the earthworm example. For example, large herbivores can graze on plants and their dung can replace the role of (higher quality) litter for earthworms (Fig. 1C). Or earthworms can profit from this dung production because it is a higher quality resource than plant litter. Even in more general terms, such loops can be found in plant–pollinator or plant–disperser interactions that also tend to be compartmentalized (like nutrient cycling; Vázquez et al. 2009). Animals profit from some form of nutrition and, in return, provide services other than nutrient recycling (e.g., fertilization, dispersal or recruitment of their host plants).

Although the universality of circular interaction structures and its importance for ecosystem organization is well accepted (as the concept of nutrient cycling), the autocatalytic (self-enforcing) nature of such loops is not. The important question then arises as to what drives the emergence of autocatalysis (causes) in these loops and how do we observe it (consequences)?

The emergence of autocatalysis

The emergence of autocatalysis in material cycles lies at the heart of resource competition theory, a link worked out in more detail by Loreau (1998, 2010a). Resource competition theory states that when multiple species are limited by the same resource, the species with the lowest steady-state resource availability ($R^*$) eventually outcompetes all other species (Tilman 1982, 1988). A lower $R^*$ can arise from higher per capita resource uptake or from lower per capita resource losses (Huisman 1994). Therefore, competition between species within a trophic level leads to selection for species with more effective resource use, which is predicted by Lotka’s maximum power principle (Lotka 1922, Odum 1971). This happens simultaneously at both the producer and decomposer side of the material cycle (Fig. 2A). In a closed ecosystem, this results in increased nutrient cycling efficiency and internal locking of nutrients over successional

Fig. 1. (A) The most basic autocatalytic loop present in all ecosystems where autotrophs and decomposing microbes form a circular configuration. (B) An extension of A including an earthworm species. A non-trophic positive feedback where earthworms improve conditions for microbes through bioturbation (blue arrow). (C) Further extension of B with the insertion of an herbivore species. Herbivores can increased the productivity of leaves through compensatory growth (blue arrow).
and evolutionary times (Fig. 2B; Loreau 1998, 2010b). Importantly, resource competition and natural selection within trophic levels result in an emergent property of autocatalysis at the level of the material cycle. Such material cycles where sets of interacting species increasingly “draw more energy and materials toward them” have therefore been termed “indirect mutualisms” as every change in primary productivity will increase secondary productivity and vice versa (Ulanowicz 2009a, Loreau 2010b). Functional traits are then expected to evolve where species promote their own conditions through other species (Barot et al. 2014). This is similar to the evolution of two-species mutualisms, but with potentially several more species involved. Here, we argue that the resulting positive feedbacks form the foundation of ecosystem organization, as they lead to the emergence of autocatalytic sets of biological species.

The ecosystem consequences of autocatalysis

In a new habitat, when plants, macrodetritivores, and bacteria simultaneously arrive (e.g., at the start of primary succession in a terrestrial ecosystem), the populations of all species are expected to grow and increasingly recycle more nutrients together, resulting in an increase in the weight of the autocatalytic loop. This increased productivity, biomass, and nutrient cycling have been shown to continue during ecological succession, when the initial plant and decomposer species become replaced by other species. This is known as ecosystem development (Fig. 2B; Odum 1969, 1983, Begon et al. 1990), although eventually a decline phase may follow (Wardle et al. 2004b). Such initial ecosystem development follows naturally from the previously described competition theory, where species are continuously replaced by new species with a higher resource-use efficiency during ecological succession. This has been shown empirically (Wedin and Tilman 1993) and theoretically (Loreau 2010b). Similarly, evolution by means of natural and sexual selection organizes ecosystems in such a way that productivity tends to increase (Leigh and Vermeij 2002). Such enhanced nutrient cycling implies strong positive feedback links within the autocatalytic loop that facilitates the capture and localization of more nutrients and energy (Bianchi et al. 1989, Stone and Weisburd 1992). Effective recycling traps nutrients into a loop and make it possible to reutilize them repeatedly, resulting in a more and more closed nutrient cycle over time. This increases the productivity of all compartments (Odum 1969, Schrama et al. 2013). This also works outside of terrestrial ecological succession. For example, the very high production of coral reefs is maintained by this type of efficient recycling of nutrients to stimulate phytoplankton growth, pelagic grazing, and bacterial (re)mineralization (Furnas et al. 2005). This effective recycling prevents leakage of internal nutrient sources and captures external source nutrients into the ecosystem. This increases the total ecosystem nutrient stock and cycling, and results in nutrient hotspots in otherwise oligotrophic landscapes. Similar insights are achieved in studies on consumer-mediated nutrient supply, for example, on promotion by herbivores of their own food availability (McNaughton 1979, Allgeier et al. 2013). Ulanowicz (2009b) refers to these cumulative nutrient retention mechanisms as “centripetality,” where the autocatalytic nature of the species interactions “pulls” more and more nutrients into a set of interacting species or locations, which in turn facilitates increased capturing of energy by primary producers. This in turn promotes energy availability to heterotrophs and enhances nutrient recycling.

Key role for ecosystem engineers in autocatalytic loops

Although consumer–resource interactions form the backbone of autocatalytic loops, species within an autocatalytic loop can potentially interact in many different ways (Olff et al. 2009), thereby changing environmental conditions for other species within the loop. Species can make local abiotic conditions for their own population growth more beneficial by habitat or niche construction, such as dam construction by beavers and soil porosity promotion by earthworms. Not only the original species profit, but also other species with associated ecological requirements (Jones et al. 1994). As a result, these positive feedbacks among species in an autocatalytic loop in combination with beneficial environmental
modification by species can shift the range of local conditions to the benefit of all participants (Ulanowicz 1997, Olff et al. 2009). Continuing with our previous example, earthworms not only increase nutrient mineralization, but also increase water infiltration, soil water-holding capacity, and aeration through their burrowing activities, which benefits both microbes and grasses (Fig. 1; blue arrows). Therefore, organisms directly or indirectly modulating the availability of resources to other species by causing state changes in biotic or abiotic material are often referred to as keystone species (Power et al. 1996), ecosystem engineers (Jones et al. 1994) or foundation species (Whitham et al. 2006). Such species are expected to fulfill prominent roles within autocatalytic loops. These environmental modifications can subsequently feedback causing a change in species composition and even evolutionary processes, i.e., via alternative niche constructions (Odling-Smee et al. 2003, Post and Palkovacs 2009, Kylafis and Loreau 2011). As Post and Palkovacs (2009) point out, such niche construction is not limited to the active engineering of the environment but includes all of the by-products of living organisms (eating, excreting, dying, nutrient uptake and mineralization, etc.). During this process of engineering the structural properties and behaviors arise at the ecosystems level through self-organization (Levin 1998).

In summary, a circular species interaction structure emerges from the simple fact that organisms frequently depend on each other for resources and evolve to benefit from each other’s waste products and environmental impacts (between trophic levels). These loops can then become autocatalytic because species compete with each for resources (within trophic levels) leading to replacement of species by others that contribute more benefits to themselves and others. This, in turn, enhances the cycling of nutrients and flow of energy at the ecosystem level. In the next section, we outline how such autocatalytic loops in ecosystems can be modeled and highlight important insights from such models with respect to ecosystem-level emergent properties.

**Modeling basic ecological autocatalysis between plants and decomposers**

Almost all ecosystems are characterized by a material cycle that needs to involve at least two key partners: (1) plants (or other autotrophs), which capture energy and inorganic nutrients to produce first living and then dead organic matter, and (2) heterotrophic decomposers, which consume this dead organic matter and release nutrients in inorganic form, to be used again by the primary producers. Their reciprocal interaction is indirect because it is mediated by the abiotic pools of dead organic matter (litter) and inorganic nutrient. The two species groups thus produce each other’s resources, making the system autocatalytic. Without either group, the resources for either plants or decomposers quickly run out, causing the system to cease to persist.

To explore the nature and functional consequences of the material cycle that results from this basic interaction, consider the simple ecosystem model depicted in Fig. 1A. The ecosystem is assumed to be limited by a single nutrient; accordingly, all compartment sizes and fluxes correspond to the stocks and fluxes of that nutrient. The inorganic nutrient pool (of size \(N\)) is supplied by a constant independent input \(I\) of inorganic nutrient per unit time. Plants produce litter, of which only that part (with nutrient stock \(M\)) readily accessible to decomposers is represented in the model (recalcitrant organic matter such as wood is not represented). Plants and decomposers have nutrient stocks \(P\) and \(D\), respectively. Their resource uptake depends on their respective stocks and functional responses to resource availability, which are represented by the functions \(f_P(N)\) and \(f_D(M)\). Plants and decomposers release nutrients as a result of metabolic processes and mortality at rates \(m_P\) and \(m_D\) per unit time, respectively. A fraction \(\lambda_P\) or \(\lambda_D\) of these flows is lost from the system, the rest \((1 - \lambda_P)\) or \((1 - \lambda_D)\) being recycled within the system in the form of readily available dead organic matter (e.g., amino acids) or inorganic nutrient (e.g., ammonia). It should be noted that \(\lambda_P\) and \(\lambda_D\) reflect the total losses from the compartments \(P\) and \(D\), respectively, irrespective of their species composition. Nutrients are also lost from the pools of inorganic nutrient and litter at rates \(q_P\) and \(q_M\) per unit time, respectively, for example by stream flow, wind, or leaching from the soil.

This model translates into the following dynamical equations (Loreau 1998):

\[
\begin{align*}
\frac{dN}{dt} &= I - q_N N - f_P(N)P + (1 - \lambda_P)m_D D, \\
\frac{dP}{dt} &= f_P(N)P - m_P P, \\
\frac{dM}{dt} &= (1 - \lambda_P)m_P P - q_M M - f_D(M) D, \\
\frac{dD}{dt} &= f_D(M) D - m_D D.
\end{align*}
\]

(1)

Assuming that primary production, \(\Phi_P\) and secondary production, \(\Phi_D\), are proportional to the nutrient inflows to the plant and decomposer compartments, respectively, we obtain, at equilibrium (denoted by an asterisk)

\[
\begin{align*}
\Phi^*_P &= f_P(N^*)P^* = \frac{S^*_N}{\Lambda}, \\
\Phi^*_D &= f_D(M^*)D^* = \frac{S^*_M}{\Lambda},
\end{align*}
\]

(2)

where

\[
\begin{align*}
S^*_N &= I - q_N N^* - (1 - \lambda_D)q_M M^*, \\
S^*_M &= (1 - \lambda_P)(I - q_N N^*) - q_M M^*, \\
\Lambda &= \lambda_P + (1 - \lambda_P)\lambda_D.
\end{align*}
\]

(3)

\(S^*_N\) and \(S^*_D\) are the net supply rates of nutrient in inorganic form and in the form of dead organic matter, respectively, at equilibrium, while \(\Lambda\) represents the fraction of the nutrient lost from the living compartments over a complete material cycle.

A key insight from this model is that the equilibrium productions of the living compartments (Eq. 2) are strongly coupled: the material cycle binds them together in a single ecosystem-level autocatalytic loop. Anything that affects one component of an ecosystem simultaneously affects all the other components of that ecosystem. Thus, material...
cycling generates an indirect mutualism between plants and decomposers, the two partners in the autocatalytic loop shown in Fig. 1A.

This simple indirect mutualistic interaction between plants and decomposers has important ecosystem-level consequences when more than one plant species and more than one decomposer species are present (Fig. 2A). Such multiple species within a functional group are competing for resources, but also indirectly benefitting both themselves and their competitors. Two sets of critical parameters are under the control of species traits and affect equilibrium productions: the resource competitive abilities; and nutrient cycling efficiencies of the various plants and decomposers. Resource competition theory predicts that the competitive ability of either plants or decomposers is determined by their ability to deplete their respective resources in a monoculture (Tilman 1982). The plant species with the lowest \( N^* \) will displace all other plant species; similarly, the decomposer species with the lowest \( M^* \) will displace all the others. Thus, competitive ability may be measured conveniently by the inverse of \( N^* \) or \( M^* \). As the competitive ability of either the plants or decomposers present in the ecosystem increases as a result of competitive replacement, the nutrient losses from the abiotic compartment they control decrease, and hence, by Eq. 3, the corresponding net nutrient supplies increase. As a consequence, ecosystem cycling efficiency and, hence, primary production and secondary production, increase (Fig. 2B). This rule applies to both ecological and evolutionary time scales, and applies to competition between types within a homogeneous material cycle, i.e., to individual selection. Thus, within-cycle competition (between different plants or decomposers) is a force that spontaneously leads to more materially closed and more productive ecosystems. In this case, evolution of these ecosystem properties (as the primary productivity of all producers together) is a by-product of the community dynamics and evolution of individual organisms. This forms a fundamental link between community and ecosystem ecology and is a key point often missed in classic food web studies that ignore nutrient cycling. It also sheds new light on classic discussions of the relative importance of competition, predation, and coevolution as evolutionary forces (Vermeij 1994). Competition within trophic levels can lead to coevolution across trophic levels and vice versa.

Plant biomass can at times also influence decomposition rates (e.g., when shading alters the soil microenvironment). In this case, the mortality rate and birth rate of the decomposers would also be functions of \( P \), not just \( M \). Such effects are generally captured in the idea of ecosystem engineering (Jones et al. 1994), which therefore can also be included in the present framework. This approach can be further extended by including competition–colonization trade-offs (Gravel et al. 2010).

Eq. 2 also predicts that species traits that improve the nutrient cycling efficiency of either plants or decomposers (i.e., that decrease either \( \lambda_P \) or \( \lambda_D \)) should have a high positive effect on ecosystem cycling efficiency, primary production, and secondary production (Fig. 2B). This nutrient conservation efficiency, however, is a trait that is selectively neutral within a homogeneous material cycle: it does not affect resource competitive ability of either plants or decomposers. Although nutrient conservation is a strategy that is strongly beneficial to all ecosystem components, the individual plants or decomposers expressing this trait would not derive any fitness benefit from it relative to their competitors. Therefore, individual selection cannot select for such traits. But selection at the ecosystem level (so for sets of species with these traits) can potentially act on these traits. For example, Berendse and Scheffer (2009) suggest this type of evolutionary mechanism for explaining how angiosperms overtook gymnosperms during the Cretaceous Period. In addition, nutrient conservation (how many nutrients are retained in the cycle) should not be confused with nutrient use efficiency. The latter is often inferred from the nutrient loss rate \( (m_P \) or \( m_D \)), which is an important component of the competitive ability for nutrients of species.

When different sets of species each are subject to ecological autocatalysis, they can start interacting, with important implications for ecosystem structure and dynamics. This is explored in the next section.

**Interactions Between Multiple Autocatalytic Loops**

*Alternative autocatalytic loops*

The presence of multiple loops (more closed as for nutrients) or channels (more open, as for energy) seems to be the rule in ecosystems rather than the exception (Baird and Ulanowicz 1989), and can be identified using algorithms (Ulanowicz 1983). Examples include coexisting “green” (aboveground, herbivore-based) and “brown” (belowground, detritivore-based) loops in ecosystems or belowground fast (bacterial dominated) and slow (fungal dominated) energy channels, respectively (Odum 1969, Moore et al. 2004, Rooney et al. 2006, 2008). The mass distribution between these different pathways has been shown to be often asymmetric (Rooney et al. 2006), with one pathway dominating over the other(s), and this dominance varying with environmental conditions across space and time (Berg et al. 2001, Neutel et al. 2007, Schrama et al. 2012). This asymmetry suggests the possibility of competition for resources (energy and nutrients) between pathways (so between sets of interacting species) with an important effect on ecosystem structure, dynamics, and functioning (Cebrian and Lartigue 2004, Bardgett et al. 2005, Butler et al. 2008, Schrama et al. 2012).

The framework of ecological autocatalysis as outlined here provides a step forward in understanding the formation and existence of such multiple autocatalytic loops. We pose two new hypotheses for the formation of multiple autocatalytic loops in ecosystems: (1) the pre-emptive resource competition hypothesis and (2) the resource diversity hypothesis.

**The pre-emptive competition hypothesis**

We suggest that a first mechanism leading to multiple autocatalytic loops in ecosystem assembly is a process of sequential pre-emptive resource competition (Fig. 3). Imagine a plant species that produces leaves that over time become litter that is subsequently mixed with soil and then decomposed by microbes (Fig. 3A). Such a loop can be invaded by a loop including macrodetritivores that get access to the resources directly after leaves fall down and become litter. The energy that the macrodetritivores respire thus pre-empts resources...
that otherwise would be used by bacteria. Along similar lines, herbivores can consume the resources as green leaves still attached to the plant before they can turn to litter, pre-empting resources that would otherwise be available to macrodetrivores and bacteria. Sequential pre-emptive resource competition can lead to alternative loops. Switches from a microb-dominated loop to a macrodetrivore- (Steinberg et al. 1997) or herbivore-dominated loop (McNaughton et al. 1997, Belovsky and Slade 2000) generally contribute to enhance nutrient cycling rates (Loreau 1995) or nutrient conservation (de Mazancourt et al. 1998). Furthermore, this order in increasingly earlier access to material produced by plants (microbes, macrodetrivores, herbivores) follows the evolution of these functional groups (Labandeira 1998, Labandeira and Currano 2013), suggesting evolutionary drivers for earlier resource exploitation, resulting in increased nutrient conservation. Similar patterns might be found at the plant nutrient uptake side. Mutualisms with specialized mycorrhizal fungi enable the uptake of nutrients as more complex organic molecules, shortcutting parts of the decomposition pathway (Hodge et al. 2001). Also, resources produced by plants that are not directly related to nutrient cycling can be subject to similar pre-emptive resource competition interactions among consumers. For example, neotropical palm trees produce fruits that are first available to arboreal frugivores (including large-bodied birds and some of the Pleistocene megafauna), then to scatter-hoarding rodents, and the remainder is targeted by invertebrates (Jansen et al. 2012). Seed dispersal distances, which are generally assumed to represent survival probability (Janzen 1970, Nathan and Muller-Landau 2000, Jansen et al. 2014), decrease from megafauna/birds to rodents to invertebrates, again suggesting that plants benefit most from organisms that can get to the resources earliest. Also here, recent evidence suggests that seed-dispersing birds evolved later than

FIG. 3. The pre-emptive resource competition hypothesis. (A) Representation of a change in plant resource from leaves to litter to litter mixed with soil with specialized consumers and associated loop. (B) Ecological (blue) and evolutionary (red) feedbacks from consumers on resources can result in different plant strategies that benefit the overall flow into each loop. The notation \( t+1 \) indicates one time step.
seed-dispersing rodents (Eriksson 2016) and those again later than invertebrate seed predators.

These processes can be reinforced when consumers not only indirectly (through nutrient cycling pathways), but also directly affect the plant resource (Fig. 3B) through ecological (blue arrows) and/or evolutionary (red arrows) feedbacks. Large herbivores, for example, can increase the productivity and nutritional quality of their forage through defoliation that keeps plants in a physiologically young active stage (McNaughton 1976, Hik and Jefferies 1990, McNaughton et al. 1997, Ruess et al. 1997). Also, grazing lawns consist of distinct grass species, with specific functional traits, that reveal a long coevolutionary history of grasses and large mammalian grazers (McNaughton 1984, Hempson et al. 2014). At first, this seems counterintuitive because defoliation should be disadvantageous to plants. However, plants adapted to grazing can get a relative fitness advantage over plants without such adaptations. Indeed, theoretical models have shown that grazing tolerance can be adaptive both ecologically and evolutionary under specific conditions (Loreau 1995, de Mazancourt et al. 1998). Furthermore, empirical data show that grazing almost always increases the nutritional quality of the plant (Hempson et al. 2014) and under specific conditions also increases its productivity. The overall result is that some plants are ecologically and/or evolutionarily adapted to grazing and attract grazers so that most (if not all) biomass produced enters the herbivore loop (Fig. 3B). Other plants might display an herbivore avoidance strategy, for example, through structural or chemical defense, and therefore most biomass produced will flow into the macrodetritivore loop. Altogether, this yields the hypothesis that the emergence of different consumer groups (and associated loops) is driven by selection for consumers that utilize plant resources earlier during their formation, and produces the emergent property of increased autocatalysis. Subsequently, ecological and evolutionary feedback of consumers on plants have produced different plant strategies that further optimize and determine the energy and material flow through each loop.

The resource diversity hypothesis

The first autotrophs in evolutionary history, as cyanobacteria, were unicellular organisms that performed all their specialized functions in a single cell and reproduced through cell division. Three billion years later, modern multicellular plants have evolved specialized organs for all different aspects of their functioning: roots for the uptake of nutrients and water, leaves for photosynthesis, seeds for reproduction, stems for structure, etc. Each of these specialized organs requires resource investment and the allocation of resources among different organs. Their structure differs greatly between species, representing a set of distinct plant strategies (Grime 1979, Westoby 1998). Such different resources produced by plants can subsequently be consumed by different groups of consumers (which is frequently intended by the plant), forming a starting point for alternative loops (Fig. 4A). Often, the consumers benefit from the resources and provide specialized services to the plant (e.g., nutrient recycling [herbivores], pollination [nectar], and seed dispersal [fruit]). Therefore, the higher the diversity in the number of resources produced by the plant, the more possible loops exist.

Feedback by consumers on the production of these resources is evident and suggests ecological autocatalysis. For example, granivorous harvester ants have been shown to increase both the number and size of the plant species involved (Rissing 1986). Furthermore, the large diversity in plant reproductive organs (flowers, fruits, seeds) suggests evolutionary feedbacks between consumer diversity and plant functional differentiation (Fig. 4B; e.g., Georgiadis et al. 1989).

The two hypotheses outlined above for the emergence of multiple autocatalytic loops in ecosystems are intricately linked. Which loop prevails or dominates in a specific situation or whether multiple loops are able to coexist depends on the interactions between loops, mediated by environmental conditions.

Alternative loops and spatial heterogeneity

As outlined before, autocatalytic sets of species may arise as a result of local internal fine-scale interactions through self-organization, and therefore provide a mechanistic basis for the occurrence of alternative energy channels. This differentiation is often a result of interacting self-reinforcing processes (Peterson 2002, Van de Koppel et al. 2002) and emerges from internal fine-scale interactions (Rohani et al. 1997). This corresponds well with interaction–redistribution models of vegetation dynamics (Lejeune et al. 2002, Rietkerk and van de Koppel 2008) based on the balance between short-range facilitation (positive feedbacks within loops) and long-range competition (negative feedbacks between loops). This is supported by empirical data on woody vegetation, with local nutrient accumulation through positive feedbacks, resulting in “islands of fertility” in nutrient-poor environments (Schlesinger et al. 1990, Belsky 1994, Callaway et al. 2002, Bruno et al. 2003). These positive interactions within loops and competition between loops may result in ecosystem-level competition between alternative autocatalytic sets of species (Ulanowicz 1997, Petchev et al. 2009).

The self-reinforcing nature of coexisting autocatalytic loops can thus be seen as a causal agent for spatial heterogeneity of landscapes at the regional scale. Fig. 5A suggests how this could work for alternative herbivore (H) and a macrodetritivore (D) loops. Plants can be consumed by large herbivores, which produce dung for dung-feeding macrodetritivores, that further process the dung to fine detritus for microbes, and that in turn mineralize detritus to provide plants with nutrients. In contrast, when plants are not eaten, they release litter fragmented by macrodetritivores (e.g., earthworms), which is further decomposed by microbes to supply nutrients for plants. Besides these consumer–resource interactions that form the backbone of both loops, they exhibit important additional feedback mechanisms. Large herbivores contribute to stress-adapted vegetation by compacting the soil, which induces hydrological and anaerobic stress (Veldhuis et al. 2014). However, macrodetritivores bioturbate the soil, increasing aeration and water-holding capacity. This generally decreases stress, with subsequent consequences for the plant community (Meysman et al.
This battle between biocompaction and bioturbation is an example of ecosystem-level competition between two autocatalytic loops that arises as a result of positive feedbacks within each loop (Howison et al. 2017). The outcome of such competition between sets of species is likely conditional on environmental conditions. In our example, rainfall is an important determinant, where plant nutritional quality decreases with rainfall and therefore becomes less attractive to large herbivores (Olff et al. 2002). Therefore, it is expected that the percentage of biomass consumed by herbivores declines along the rainfall gradient, resulting in a gradual shift from an herbivore- to a detritivore-dominated system (Fig. 5C). Importantly, the biocompaction and bioturbation mechanisms promote the dominance of their loops, by shifting the local conditions toward favorable conditions (Howison et al. 2017). Biocompaction by large herbivores creates locally dry conditions due to reduced water holding capacity and infiltration (Thurow 1991, Veldhuis et al. 2014). Bioturbation by macrodetritivores, on the other hand, increases infiltration rates and water-holding capacity (Joschko et al. 1989, Meysman et al. 2006), pulling the system to the wet end of the gradient and stimulating the detritivore loop. This results in the well-known grazing-lawn–bunch-grass mosaics at intermediate rainfall (400–1,200 mm/yr). Locally, either loop can dominate (Howison et al. 2017), while at larger scales, a gradual shift from an herbivore-dominated to a detritivore-dominated system is observed (Fig. 5B).
Modeling multiple autocatalytic loops in ecosystems

A key finding in the study of autocatalytic processes is that a significant fraction of nutrient cycling takes place at much smaller spatial and temporal scales than previously believed. For instance, about two-thirds of nitrogen uptake by grasses originates from rapid mineralization of dead roots within their rooting system in some tropical savannas (Abbadie et al. 1992). These grasses even control nitrification in their immediate vicinity through a balance between inhibitory and stimulatory effects on nitrifying bacteria (Lata et al. 2000, 2004). This is to their own benefit, as nitrate is easier to metabolize for plants than ammonia. In this case, a relatively tight association between individual plants and microbial populations should be expected. In addition, evidence is accumulating that grasses in nutrient-limited conditions can promote nitrogen fixing bacteria in their rhizosphere through exudation, which promotes directly available nitrate in the immediate vicinity of their roots (Gupta et al. 2014). Such strongly localized spatial structures tend to generate between-cycle competition, i.e., competition between sets of organisms involved in spatially distinct cycles. Such tight associations between plants and rhizosphere micro-organisms involved in decomposition can imply that plants indirectly compete with free-living microbial decomposers for detritus instead of mostly being facilitated by them. Further evaluation of the importance of such mechanisms is needed.

As an extreme case of such between-cycle competition, consider a perfectly structured environment in which each individual plant occupies an isolated site during its lifetime and is associated with a single decomposer individual or colony of similar life span. Assume that sites become vacant when previous occupant pairs are extirpated by natural death or disturbance, and establishment of both plants and decomposers at vacant sites obeys a competitive lottery. Finally, assume that the probability of a genotype’s successful establishment at a site is proportional to its total production in all other sites, because higher production means production of more propagules of a higher quality.

The dynamics of site occupancy by plants then obeys the equation (Loreau 1998)

$$\frac{dp_{Pi}}{dt} = p_{Pi}(r_{Pi}V_{Pi} - m_{Pi})$$

(4)

where

$$r_{Pi} = \alpha_{Pi} \left( \frac{\sum_{j} \Phi_{Pi}}{p_{Pi}T} \right)$$

(5a)
$V_P = 1 - \sum_k p_{Pi}^k.$  \hspace{1cm} (5b)

In this equation, $T$ is the total number of sites available, $p_{Pi}$ the proportion of sites occupied by plant genotype $i$, $m_{Pi}$ its mortality rate, $\Phi_{Pi}$ its productivity at site $j$ (see Eq. 2), and $\alpha_{Pi}$ its reproductive efficiency, a constant of proportionality that incorporates both the allocation of plant genotype $i$’s production to reproduction and its ability to disperse and establish at new sites. The aggregated parameter $r_{Pi}$, which is plant genotype $i$’s average productivity times its reproductive efficiency, represents a potential reproduction rate, reproduction here being considered completed after the establishment of offspring at new sites. Last, $V_P$ is the proportion of vacant sites; only dispersal to vacant sites leads to successful reproduction.

An equivalent equation holds for decomposers with a mere change in subscripts

$$\frac{d p_{Di}}{dt} = P_{Di}(r_{Di} V_D - m_{Di}).$$ \hspace{1cm} (6)

At equilibrium, the fraction of vacant sites, $V_X^\infty$, in Eqs. 4 and 6 must satisfy

$$V_X^\infty = \frac{m_{Xi}}{r_{Xi}}.$$ \hspace{1cm} (7)

where $X = P$ or $D$. This relation can be satisfied only by a single species or genotype. Therefore $V_X^\infty$ here plays the same role as $R^\infty$ in classical resource competition, and the species or genotype with the lowest $V_X^\infty$, and hence the highest basic reproductive capacity (the inverse of $V_X^\infty$), eventually displaces all the others.

In the simplest case, where plants and decomposers disperse independently and their effects on their local environment are additive, the outcome of this dual selective process is the selection of the material cycle that combines the plant and decomposer genotypes with the highest basic reproductive capacities. Since the basic reproductive capacity of a genotype is proportional to its average productivity at a site (Eq. 5a), traits that contribute to increasing equilibrium productivities may be selected for. This assumes that dynamics of site occupancy (dispersal, establishment) is slow compared with the dynamics of material cycles (nutrient uptake, growth) within sites. In particular, selection for increased nutrient conservation is possible, leading to enhanced ecosystem properties, in particular increased ecosystem cycling efficiency and primary and secondary productivities.

A feature of this scenario is that material cycles within sites behave very much like “superorganisms” (Wilson and Sober 1989), where genotypes play the role of alleles at the plant and decomposer “loci” and the basic reproductive capacity is the measure of fitness. Like organisms, these spatially separated cycles have a temporary existence: their properties result from the random assortment of their constituent genotypes, and the unit of selection is the entire meta-genome of the species set. Selection of traits advantageous to the whole cycle (set of interacting species) is then just as natural as selection of traits advantageous to the individual organism in classical individual selection theory.

Unlike organisms, however, the biotic components of the material cycle reproduce separately, but this does not affect the overall direction of the outcome.

This so-called ecosystem selection is a rather extreme evolutionary scenario that requires species interactions to be both long-lasting and strongly localized (Loreau 2010a). Therefore, one should generally expect ecosystem selection to co-occur with, and often be weaker than, individual selection. In Modeling basic ecological autocatalysis between plants and decomposers, we showed that classical individual selection driven by within-cycle competition can sometimes lead to similar evolutionary outcomes as ecosystem selection. But this convergence of outcomes is not expected to hold generally, in particular when there is pre-emptive competition between different autocatalytic loops.

Consider, for example, the widespread case of plant–herbivore interactions. Both plants and herbivores recycle limiting nutrients, leading to two alternative recycling pathways. But herbivores eat plants, hence there is a direct antagonism between the two partners. The ecological and evolutionary dynamics of this interaction are much more complex than in the indirect positive interaction between plants and decomposers. Over ecological time-scales, plants can benefit from the presence of herbivores in the form of enhanced plant production despite the negative direct effect of herbivores on their biomass. This occurs when herbivores recycle limiting nutrients more effectively than do plants, i.e., when they enhance the overall nutrient conservation efficiency of the ecosystem, and grazing intensity is not too high (de Mazancourt et al. 1998). Thus, the more efficient alternative autocatalytic loop provided by herbivores benefits plants indirectly. This is despite the direct cost plants incur from being eaten and generates an indirect mutualism between the two partners. Such indirect effects of predation benefiting prey are surprisingly widespread in ecosystem networks and play a much more important role than is generally assumed (Bondavalli and Ulanowicz 1999).

Over evolutionary time scales, however, this ecological benefit is not necessarily selected for. Indeed, it is not absolute, but relative fitness that counts. If two plant types are mixed, one of them being tolerant (mutualistic) and the other resistant (antagonistic) to herbivory, the resistant type is expected to outcompete the tolerant type because it benefits from the positive indirect effect of increased nutrient cycling (as promoted by the tolerant plant) but does not suffer the negative indirect effect of herbivore consumption. As a result, the fitness of the resistant type is higher than that of the tolerant type and tolerance does not evolve, even though it is indirectly beneficial to both types. Two factors can counteract the advantage of antiherbivore defense: spatial heterogeneity and the cost of defense. Just as in plant–decomposer interactions, spatial heterogeneity in nutrient recycling by herbivores can lead to the selection of “mutualistic” tolerant plant types if herbivores recycle nutrients in the vicinity of the grazed plants or if plants from the same type are aggregated (de Mazancourt and Loreau 2000). The cost of defense generates a trade-off in plants between antiherbivore defense and growth or nutrient uptake, which can lead to a complex set of evolutionary outcomes including plant–herbivore mutualism (de Mazancourt et al. 2001). The interplay between ecological and evolutionary outcomes, however, is so complex that it
required a redefinition of the very concept of mutualism (de Mazancourt et al. 2005).

Similar to the plant–herbivore example, Harte and Kinzig (1993) modeled the dynamics of microbial decomposers that compete with plants for inorganic nutrients and also benefit from plants through their carbon input via dead organic matter. Therefore, there is a direct negative effect of plants on microbes but, at the same time, an indirect positive effect. Also, the indirect mutualistic interactions could prevail only in a spatially explicit model where local autocatalysis promoted “mutualistic microbes” that were able to outcompete “selfish microbes” (Kinzig and Harte 1998). In contrast, in a homogeneous environment mutualistic microbes were not selected for. These examples of plant-microbe and plant-herbivore interactions reveal the huge potential of jointly considering ecological and evolutionary dynamics of autocatalytic loops in ecosystems.

PERSPECTIVES

Throughout this paper, we have reviewed the importance of ecological autocatalysis as a key internal driver of ecosystem organization. Furthermore, we have emphasized that the concept of autocatalytic sets of species roots in biochemistry and systems biology (Eigen and Schuster 1979, Morowitz et al. 2000, Lincoln and Joyce 2009, Giri and Jain 2012). This may hint at a basic process that re-emerges across levels of organization, and suggests generality of autocatalytic sets as a driving force of structure across all levels of biological organization (Gadgil and Kulkarni 2009). These “nested autocatalytic sets” can stabilize higher level structures, even with relatively low catalytic strength, suggesting statistical or thermodynamic favor over alternative configurations (Giri and Jain 2012). These nested autocatalytic sets now require further quantification and theoretical study, especially with regard to the interplay of ecological and evolutionary dynamics.

The universality of circular interaction structures follows from the simple fact that organisms produce resources that are subsequently used by other organisms, culminating in chains of such resource-consumer interactions. As many chemical elements are essential for life and often limiting, the reuse of such elements results in the closure of such interaction chains (formation of loops), where material circulates and autotrophs fuel these interaction structures with an input of energy. Autocatalysis emerges in such circular interaction structures through basic principles from community ecology (resource competition and natural selection) and evolutionary biology (with indirect mutualism as an extended form of coevolution). As a result, nutrient availability and nutrient cycling can be viewed as emergent properties of the actual configuration of trophic and non-trophic interactions at the ecosystem level (De Ruiter et al. 1994, Berg et al. 2001, Vos et al. 2011). This is in contrast to the accepted view that they “determine” the outcome of species interactions. In addition, development of the concept yields key patterns observed at the ecosystem scale, such as alternative stable states, landscape heterogeneity, and ecosystem resilience. These arise as a consequence of autocatalytic loops instead of having to be seen as independent processes (Fig. 6). Furthermore, the resulting system-level interaction structures (autocatalytic loops) have large consequences for community structure and evolutionary radiation instead of the classic view that community structure is only determined by functional traits or assembly rules of component species, which in turn determines ecosystem functioning (bottom-up causality).

Ecosystems have been referred to as complex adaptive systems because their macroscopic properties, such as patterns of nutrient and biomass flux, diversity–productivity relationships, or trophic structure, emerge from local, small-scale interaction where interaction structures reflect self-organization (Levin 1998, Morowitz 2002). More importantly, the resulting macroscopic properties may feed back as a selective force to lower levels of organization (diffuse feedback), and affect the future development of its components. These higher-level structures (biotic and produced abiotic conditions) are a large part of the environment of an evolving organism (Kauffman 1993, Odling-Smee et al. 2003, Hoelzer et al. 2006, Hastings et al. 2007, Matthews et al. 2014). Therefore, self-organization and natural selection should be

![Fig. 6. Overview of the positioning of the concept of ecological autocatalysis. Autocatalytic loops represent the interaction structure of species through self-organization and follow from principles in community ecology. At the same time, the autocatalytic loops feedback on the individual species as they constitute their environment. Last, autocatalytic loops form the mechanistic basis for understanding the processes and patterns observed within ecosystems.](image-url)
seen as processes that interact across levels of organization (Fig. 6). At the ecosystem level, ecological autocatalysis “creates” the environmental conditions that different species in different loops encounter. Hence, they shape the course of natural selection, which changes the role different species may play in an autocatalytic loop with system-level consequences.

Conclusion

The historic focus of ecology on pairwise interactions and on responses of species to ecological factors has obscured the importance of higher level ecosystem organization and species–environment feedback. The framework of ecological autocatalysis proposed here aims to include all these interactions and at the same time reduce overall complexity. We suggest that it provides a rich set of opportunities in further developing, formalizing, modeling, and experimentally testing the fundamental principles of ecosystem organization.

Acknowledgments

M. P. Veldhuis, M. P. Berg, and H. Olff together developed the concept. M. P. Veldhuis wrote the first draft of the manuscript. M. Loreau contributed the modeling sections and all authors contributed substantially to the revisions. M. Loreau was supported by the TULIP Laboratory of Excellence (ANR-10-LABX-41). We thank Eric S. Higgs and two anonymous reviewers for their valuable comments on earlier versions of the manuscript. This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 641918. [Correction added 12 June, 2018 after online publication. The last funding statement was mistakenly omitted from the original manuscript.]

Literature Cited


