

1 **Unifying concepts of biological function from molecules to ecosystems.**

2

3

4 **Keith D Farnsworth^{1*}, Larissa Albantakis², Tancredi Caruso¹**

5

6

7 1. Queen's University Belfast. MBC. 97 Lisburn Road, Belfast BT97BL, UK.

8 2. Department of Psychiatry, University of Wisconsin, Madison 53719, WI, USA.

9

10

11 Keith Farnsworth

12 * Corresponding author: k.farnsworth@qub.ac.uk

13 Orchid ID: 0000-0001-8025-1114

14

15 Larissa Albantakis: albantakis@wisc.edu

16

17 Tancredi Caruso: t.caruso@qub.ac.uk Orchid ID: 0000-0002-3607-9609

18

19

20

21

22 **Abstract**

23 The concept of function arises at all levels of biological study and is often loosely and
24 variously defined, especially within ecology. This has led to ambiguity, obscuring the
25 common structure that unites levels of biological organisation, from molecules to ecosystems.
26 Here we build on already successful ideas from molecular biology and complexity theory to
27 create a precise definition of biological function which spans scales of biological organisation
28 and can be quantified in the unifying currency of biomass, enabling comparisons of
29 functional effectiveness (irrespective of the specific function) across the field of ecology. We
30 give precise definitions of ecological and ecosystem function that bring clarity and precision
31 to studies of biodiversity-ecosystem function relationships and questions of ecological
32 redundancy. To illustrate the new concepts and their unifying power, we construct a simple
33 community-level model with nutrient cycling and animal-plant mutualism, emphasising the
34 importance of its network structure in determining overall functioning. This type of network
35 structure is that of an autocatalytic set of functional relationships, which also appears at
36 biochemical, cellular and organism levels of organisation, creating a nested hierarchy. This
37 enables a common and unifying concept of function to apply from molecular interaction
38 networks up to the global ecosystem.
39

40 **Function as a unifying concept for all levels of biological organisation.**

41

42 Are all the elements of a genome functional or are some junk? Are some species ecologically
43 redundant or do they all have a unique role? Must a thing be naturally selected before it can
44 be accepted as functional, or is it sufficient that it causes an effect? These questions all hinge
45 on the definition we use for function. This became one of the key topics in a recent multi-
46 disciplinary workshop “Functional Information: its potential for quantifying biodiversity and
47 its relation to ecosystem functioning”, organised by the Synthesis Centre of Biodiversity
48 Sciences, Germany. The word ‘function’ has a meaning that may initially seem self-evident
49 and obvious, but it is used loosely and for many different meanings in biology, creating
50 ambiguity and uncertainty which matters when we come to quantify e.g. biodiversity
51 ecosystem function relationships. Here we draw ideas from molecular biology, community
52 ecology, systems and information theory and philosophy of science to construct a precise and
53 quantifiable definition of biological function that can unify and focus our thinking on the
54 questions above. The context in which we write is one of a growing importance for functional
55 descriptions in ecology (Krause et al., 2014; Stouffer et al., 2012), the rise of metagenomic
56 functional analysis (Deng et al 2012; Fierer et al., 2012; Howe et al., 2014) and controversy
57 over functional elements in, especially, the human genome (Kellis et al., 2014; Doolittle et al.,
58 2014). Because ‘function’ is used differently among sub-disciplines of ecology and wider
59 biology, our aim here is to promote a unifying meaning, with three goals. First, to enable
60 cross-fertilisation of ideas through the recognition of a common concept of function, so
61 making connections and enabling the discovery of more fundamental principles. Second, to
62 enable communication by using ‘function’ terms in a precisely defined way. Third, to promote
63 function as a precise, quantitative concept (with units) to establish a common framework for
64 biodiversity-ecosystem function studies and to enable e.g. objectively based economic
65 valuation of ecosystem components and diversity (see Farnsworth et al., 2016a). More
66 deeply, we aim for a definition of ecological function that can illuminate the extent to which
67 ecosystems are coherent systems, as opposed to mere arbitrary assemblies.

68

69 We start by clearly differentiating between ‘function’ and ‘trait’ in the biological context.
70 Violle et al., (2007) dispelled much confusion over the term trait, but still today various kinds
71 of processes and even the values of indices calculated from samples are sometimes
72 considered ‘traits’ and this is confusing. Here, we restrict function to describe a process (an

73 action) rather than a property of an entity. This leaves the latter to be a trait: a trait is an
74 inherent property of a biological system (e.g. organism), which may enable a function to be
75 performed in relation to another system (e.g. a community). For example, the specific
76 behaviour of hiding certain plant seeds is a trait of some mammals which has the functions of
77 dispersing the plants as well as providing a continuous food supply for the mammal.
78 Hypothetically many more functions of this behavioural trait are conceivable, but in practice
79 we must restrict our considerations to actual (observed) and ecologically relevant functions,
80 as opposed to potential (hypothesised or not) functions because the full range of potential
81 functions is indeterminate (e.g. providing food for soil invertebrates may be admissible, but
82 providing a means of counting the mammals is not, since the latter is not an ecologically
83 relevant function, but merely a tool of scientific observation).

84
85 The specific meaning of a function depends on both the functioning entity and the context in
86 which it acts (function is relational, not inherent). For example a genetic element may code
87 for a protein only in particular circumstances and a species may promote nitrogen turnover
88 only in certain environments, therefore they only have those functions in those particular
89 contexts. The context, in every case, is the system in which the functioning entity takes part.
90 In a biological setting we refer to this as the next higher ‘ontological level’ recognising that
91 living systems are composed of a nested hierarchy of organisation: systems constructed from
92 sub-systems (Farnsworth et al., 2013). An ontological level, in this context, is a structure of
93 biological organisation that is categorically different from those above and below in the
94 hierarchy. For example, though a nuclear family and a whole tribe differ in organisational
95 scale, in category both are units of human society. Conversely a collection of living cells and
96 a human being are both different in organisational scale and different categories of entity.
97 More formally, an ontological level (in biology) has associated with it an organisational
98 structure with the properties of a complex system, especially the property that new
99 phenomena emerge within it (see also Marshall et al., 2016). An ontological level is always
100 more than a collection of objects. Thus, for example, a population (defined in the traditional
101 sense of a mere aggregate, or count, of individuals) is not ontologically distinct from the
102 individual organism because no properties emerge from the population level, but if the
103 population is considered as an assembly of genes, then it may qualify as ontologically distinct
104 from the single genome because emergent properties may appear at the level of gene-
105 population. Further, a community may be a complex system with an organising structure that

106 leads to emergent properties (e.g. trophic cascades or reaction-diffusion patterns), so we
107 categorise it as a distinct ontological level. We illustrate how life as a whole can be viewed as
108 a hierarchy of such ontological levels (Farnsworth et al., 2013) using the ontological
109 hierarchy given in Table 2, but this does not imply that those *specific* demarkations are fixed
110 or necessary for our understanding of function.

111

112 From here on, we define biological function with respect to a higher ontological level. We
113 must carefully distinguish between the terms ecological function and ecosystem function.
114 Given our axioms, ‘ecological function’ must be an act performed by a living system *within*
115 *the context of* an ecosystem. That applies equally whether the act is physiological,
116 behavioural, one of competition, predation or mass-transfer, ecosystem engineering, or
117 organisational - every branch of ecology can make use of the same concept here. Ecosystem
118 function is conversely an act performed *by an ecosystem* in some ontologically higher system
119 - but what? Many ecologists seek a justification for conservation in the ecosystem services (a
120 result of functions) provided for human society, tacitly assuming humanity to be the wider
121 context (e.g. deGroot et al., 2012). A non-anthropocentric (objective) account places things
122 the other way around (indeed this is a central tenet of ecological, as opposed to
123 environmental, economics (Costanza and Daly, 1987). Objectively, the higher ontological
124 level of ecosystem function must comprise the global geo-chemical cycle that includes the
125 biosphere performing certain homeostatic processes (Lovelock and Margulis, 1974).

126

127 In his landmark paper, Jax (2005) resolved the use of the word ‘function’ in ecology into four
128 broad meanings: 1) individual-level processes, such as a particular predation event; 2)
129 systemic processes, such as nitrogen uptake; 3) individual ‘roles’ defining ‘functional groups’
130 of organisms, including their contribution to a higher level of organisation, such as the guild
131 of detritivores, or simply a phenotypic (often life-history) category and 4) effects of the
132 activity of working ecosystems that impinge on human society, leading to ecosystem services.
133 We see immediately that (3) describes a set of functional capabilities of organisms and (1)
134 describes the realisation of one or more of these in practice, whilst (2) does the same but
135 takes a coarse-grained view of the system and (4) interprets this from an anthropocentric
136 perspective.

137

138 In the following, using examples from ecology and molecular biology, we will argue for a

139 more constrained definition of function, in line with (3): that a biological function must
140 demonstrate **causal effect** (Table 1) from one to a higher ontological level of biological
141 organisation. This relationship enables us to identify ‘**functional equivalence sets**’ (Table 1),
142 shifting the focus from particular systems, defined by their traits, to functional classes. This
143 move enables a logically self-consistent and integrative way to assess all kinds of interaction
144 in biological systems, by starting from a network of functional equivalence sets. Only by
145 doing this can we be sure that function has a standard and inter-comparable meaning, in for
146 example biodiversity-function relationships.

147

148 **Insights from philosophy and molecular biology**

149

150 A system composed of parts, each of whose existence depends on that of the whole system is
151 termed a ‘Kantian whole’ (Table 1), the archetypal example being a bacterial cell (Kauffman
152 and Clayton 2006). The concept of function arises to describe the role that the parts play in
153 the system, without having to descend into non-scientific teleology. (Note - the origin of this
154 terminology lies in Emanuel Kant’s definition of an organised whole (Ginsborg, 2006)).
155 Accordingly, the philosopher Cummins (1975) proposed that ‘function’ is an objective
156 account of the contribution made by a system’s component to the ‘capacity’ of the whole
157 system. Thus *Cummins function* (Table 1) describes a relationship between a system and one
158 or more of its component parts. The relationship is that at least one process performed by the
159 component/s is *necessary* for a process performed by the whole system, as in the case of
160 category (3) of the Jax (2005) schema. This is a powerful idea we shall use from here on.
161 What we add is clarification about what specifies the ‘whole system’, particularly how it is
162 bounded, and how this concept can be turned into a tool for quantifying function, especially,
163 expanding the notion of being ‘necessary’ to a rigorous quantification of causal effect.
164 Natural selection for a particular component within an organism provides a clear example. In
165 that case, the component (e.g. an eye) contributes a process that increases the (quantifiable)
166 Darwinian fitness of the organism, thereby demonstrating Cummins function. The idea, that
167 ‘true biological functions’ should be naturally selected, has already emerged in the field of
168 molecular biology, though a broader set of function is advocated by some - as we next
169 discuss.

170

171 Recently, thousands of full genomes have been sequenced, not the least being the human

172 genome, and efforts to interpret this new data have stirred up the controversy over ‘junk
173 DNA’, which is taken to be the alternative to ‘functional DNA’ (see [Kellis et al., 2014](#)). In
174 this field, little attention has been paid to the precise meaning of ‘function’, despite much
175 loose talk of ‘functional DNA’. In their criticism, [Doolittle et al., \(2014\)](#) clarified the range of
176 potential meanings of the term and the conceptual errors that may arise from failing to
177 resolve them. The most justifiable assignment of function to a DNA element, they argued,
178 was where selection at the organism level was demonstrated. This is only a subset of cases
179 where a causal role has been established. Alternatives include a) selection at higher or lower
180 ontological levels, b) neutral (non-selective) processes, which frequently ratchet their way
181 into conserved stability, and c) spandrels (the term used by [Gould and Lewontin \(1979\)](#), for
182 by-products of selection). [Doolittle et al. \(2014\)](#) were careful to distinguish the spandrels from
183 “mere effects”, which play no causal role in the system to which they belong: spandrels play
184 an unintended and perhaps irrelevant causal role (e.g. the ‘thumping sound’ of the heart).

185

186 Building on this, we may at least ascribe the word ‘functional’ to biological systems having
187 causal effects that are known to have been naturally selected for at any level of organisation
188 subject to natural selection (not only species, but also gene, or genetic network). They are
189 functional in the Cummins sense and their causal effect can be quantified by the change in
190 fitness they cause in the system to which they belong. However, this definition would leave
191 out systems with Cummins function but not derived from natural selection. They would be
192 left as ‘merely’ causal, rather than biologically functional. These causal systems could be
193 subsystems of naturally selected systems, for example ‘molecular machines’ (e.g. the ATP
194 synthase complex), or biochemical networks (examples in [Jaeger and Calkins, 2012](#)), or they
195 may be super-systems of them: populations or communities of organisms.

196

197 Metagenomic functional analysis challenges this restrictive notion of biological function by
198 revealing the processes that are being (RNA), or could be (DNA) performed at the
199 community level of organisation (Tringe et al., 2005; Warnecke et al., 2007; Huttenhower et
200 al., 2012; Howe et al., 2014). With it, function is readily discerned because the link between
201 gene expression and functional protein or peptide is usually known and unambiguous,
202 irrespective of our knowledge of its natural selection. Although eukaryotic organisms
203 typically present more complicated and often less clear geno-phenotypic links, the method is
204 also developing for them (see e.g. Knack et al., 2015). These new molecular tools reveal a
205 literal functioning: the causal chain leading from one biomolecule to the next. The
206 components (e.g. functional genes) are performing Cummins function for the organism level
207 (which because of this, performs Cummins function at the community level) and their
208 performance is well defined and quantifiable. This unambiguous identification of a causal
209 chain is the ideal towards which ecologists may strive (Gotelli et al., 2012; Bohmann et al.,
210 2014). Accordingly, we now offer a definition for biological function:

211
212 **A biological function is a process enacted by a biological system A at ontological level n**
213 **which influences one or more processes of a system B at level $n+1$, of which A is a**
214 **component part.**

215
216 This definition is similar to that of Cummins function, but more precisely identifies the link
217 between ontological levels of organisation. By tracing functional effects through the nested
218 hierarchy leading from one to the next higher ontological level, it can be made explicit how a
219 function at the molecular level can be functional at an ecological level. The definition does
220 not require a history of natural selection and it does not require a 'good' outcome for the
221 system (as selection would). The latter point answers the philosophical objection often
222 levelled at scientific accounts of function (see e.g. Griffiths, 1993), which arises from the
223 teleological (referring to purpose) and normative (what ought to be) connotations of the
224 every-day use of the word 'function' (a rich literature on philosophical approaches to
225 biological function is reviewed by Neander (2011)). Our definition creates a clear separation
226 between these everyday meanings of function and a strict scientific meaning.

227
228 If functions are strictly processes, then potentially more than one system component can

229 perform them and functional redundancy and substitution become possible. This idea, which
230 we now incorporate into our definition of function is usefully formalised by the concept of
231 *functional equivalence class* (FEC) (Table 1), which has grown from analysis of biochemical
232 networks. The FEC was defined by Auletta et al., (2008) as a set of biochemical ‘operations’
233 having effects in common which are relevant to ‘goals’. The FEC consists of all operations
234 (behaviours or processes) having the effect in question and this is context-dependent because
235 an effect always depends on the nature of both the subject and the object. For example, the
236 DNA sequences and corresponding protein structures of alcohol dehydrogenases in
237 vertebrates bear no similarity with those of *Drosophila* and they work through different
238 chemical reactions, but achieve the same end result of removing hydrogen from alcohol
239 (Doolittle 1994). Those different dehydrogenating processes form an FEC with respect to
240 frugivorous organisms (and would not with respect to obligate carnivores). The existence of
241 alcohols in fruit arises from interactions among plants and micro-organisms (i.e. at the
242 community level) and exerts natural selection on the metabolic processes (cellular level) of
243 frugivorous organisms. This illustrates a general feature of FECs: they define the function at
244 one ontological level in terms of processes at a higher level, which in turn must include at
245 least one component from among them. Any function specifying an FEC complies with the
246 definition of Cummins function. Using the FEC generalises the concept by identifying the
247 source of causation: it is not the individual components performing the function, but the FEC
248 (i.e. the macro-level unit). This allows for the possibility of functional redundancy and
249 substitution: both important features illustrated at the ecological level by species turn-over
250 and at the sub-species level by heterologies and convergence (reviewed by McGhee, 2011).
251 We next discuss the extent to which this definition of biological function applies in ecology,
252 with special reference to the biodiversity-function relationship.

253

254

255 **At what level do whole ecosystems function?**

256

257 Function in ecological communities had traditionally been thought different from function in
258 whole or parts of organisms because ecological systems do not fit the conventional model of
259 Darwinian evolution (Maclaurin and Sterelny, 2008). Recent evidence (e.g. Rillig, et al.,
260 2015) and an emerging ecological theory of broader evolution (Laland et al., 2015)

261 challenges that view. Niche construction theory (Odling-Smee et al. 1996) and the concept
262 of reciprocal causation in which organisms influence their evolution via ecological processes
263 (Laland et al., 2011) both include the ecosystem as an integral part of a complex evolutionary
264 process.

265

266 The predominance in community ecology of predator-prey networks, described by flows of
267 energy in foodwebs has emphasised an incomplete model of community structure in which
268 only one function (energy flow) is admitted, rendering what Loreau (2010) calls ‘horizontal
269 diversity’ redundant. In every real community, each organism performs more than one
270 function and functions are more than contributions to energy flow. This has been recognised
271 in the development of community models based on nutrient-cycling (e.g. Loreau, 1996;
272 Thébault and Loreau (2003); Loreau (2010) and in the study of mutualistic networks
273 involving pollinators and seed dispersal (e.g. Schleuning et al., 2015), to which parasites and
274 ‘ecosystem engineers’ might be added (see e.g. Bruno et al., 2003). As Krause et al., (2014)
275 explicitly state, ecosystem functioning results from the interactions among organisms, which
276 though based on organism diversity are a (neglected) form of diversity in themselves. Our
277 definition of function as strictly relational calls attention to the importance of the interaction
278 network as a component of diversity and it identifies the organisation (network structure) of
279 its component ecological functions as the (proximate) cause of ecosystem function. This is a
280 point we wish to emphasise, especially because so far, the mainstream of studies into
281 biodiversity / ecosystem function (BEF) relationships has neglected the contribution of the
282 community network (in and of itself) to ecological functioning (Hooper et al., 2005;
283 Cardinale et al., 2012), despite contradictory evidence (e.g. Bascompte, 2009; Grey et al.,
284 2014; Valiente-Banuet et al., 2015). Experimental manipulations that discriminate the effect
285 of network structure (the organisation of functions at the community level) from organism
286 effects (the traits enabling ecological functions) will be challenging, but are necessary to
287 appreciate the extent to which the higher ontological level of community operates.

288

289 **What this says about functional redundancy**

290

291 Functional redundancy does not mean that a system will be unaffected by species loss. The
292 idea that species may be mutually redundant can arise from the neglect of the quantitative
293 contributions made by *individual* organisms to ecological processes. Functionally

294 substitutability (belonging to the same FEC) does not imply mutual redundancy: an analogy
295 with members of a tug-of-war team illustrates how individual organisms can be functionally
296 substitutable, but not redundant. In community ecology, this point is well illustrated by
297 empirical findings, e.g. from O’Gorman and Emmerson (2009) and Isbell et al., (2011).
298 Further, we must distinguish between redundancy and degeneracy. Consider a set of n species
299 \underline{S} , and a set of k functions \underline{F} . We associate a k -long vector of functions \underline{f}_i with each member
300 species s_i , the values in the vector being the quantified contributions of s_i to each of the k
301 functions in \underline{F} . Now we can transform from n species to m functional equivalence sets \underline{E} by
302 gathering all those members of \underline{S} sharing the same \underline{f} . But being relational, functional
303 equivalence must be defined relative to the context, i.e. the environmental conditions.
304 Organisms that are members of FEC e_j in \underline{E} , irrespective of environment are qualitatively
305 redundant, but those that only share \underline{f}_j in common for a particular environment are merely
306 degenerate with respect to that environment alone (Tononi et. al., 1999). This is the basis for
307 the ‘insurance’ justification for biodiversity in the environmental economics literature (Yachi
308 and Loreau, 1999; Baumgärtner, 2007).

309
310 Community ecology and especially BEF research has tended to account for one single
311 species-level Cummins function at a time (e.g. total energy flow), though this constriction
312 was eased, notably by Gamfeldt et al., (2008) and Isbell et al., (2011) and by attempts to
313 integrate mutualistic networks along with competition, predator prey and parasitism (see Kefi
314 et al., 2012 and reviews by Krause et al., 2014 and Namba, 2015). Since each organism can
315 perform many functions and each function can be performed by many organisms, it makes
316 sense to transform from networks of species-specific populations to quantitative functional
317 networks of FECs. Many community ecologists have started to refocus this way (e.g.
318 Schleuning et al., 2015), aided by increased data from molecular ecology which can more
319 precisely isolate function (e.g. Zhou et al., 2010; Deng et al., 2012; Fierer et al., 2012). There
320 is insufficient evidence to claim that any organism is functionally redundant in the
321 quantitative, multi-functional sense (Gamfeldt et al., 2008, Farnsworth et al., 2012) and we
322 think it unlikely.

323

324

325 **Quantifying function**

326

327 Identifying FECs is a key component of the functional analysis of biological systems.

328 FECs may be equated with observed functions from metagenomic functional analysis (see
329 e.g. [Fierer et al., 2012](#)) and other identified processes such as metabolism, predation and
330 specific ecosystem engineering. In practice, this may result in too many classes for
331 quantitative analysis. We would then be forced (as we usually are) to identify a set of *a-priori*
332 function classes to which observed functions are assigned. Choosing these classes will
333 depend on what degree of difference among functions we would consider sufficient to
334 identify them as separate. This may be objectively estimated from the overlapping of the
335 functions' effects, for which the concept of 'effective number' in biodiversity ([Jost, 2006](#),
336 [Chao et al., 2014](#)) was applied to ecosystem processes by [Ulanowicz et al., \(2014\)](#). Observed
337 functions would be assigned in proportion to both the number of components (enzymes,
338 populations, etc.) and the effectiveness of the functions in any given environment.
339 Effectiveness is empirically quantifiable given a precise definition of the function as a
340 process. For some processes this approach is already well established. For example, predation
341 is quantified by the individual (per-capita) 'functional response' (see e.g. [Dick et al., \(2014\)](#)
342 for a relevant application). Ecosystem engineering is often highly context specific, but
343 empirical approaches have been found (e.g. [Queiros et al., 2011](#)). The approach we are
344 suggesting calls for a precise definition and identification of FECs, for which the payoff is
345 likely to be a more robust understanding of how biological systems work.

346

347 As we have defined them, functions are causal. Thus in principle we can quantify them in
348 both magnitude and direction, according to their causal effect (for which structural equation
349 modelling and related methods may be useful ([Shipley, 2000](#)). In practice this can be very
350 difficult for complex systems with reciprocal causation ([Laland et al., 2011](#)), particularly in
351 the face of limited time-series data samples, latent variables, and incomplete knowledge of
352 the system under consideration (see [Sugihara et al., 2012](#) and [Deyle et al., 2016](#)). Recent
353 developments in information-theoretic and network analysis (e.g. [Zenil et al., 2016](#))
354 combined with knock-out experiments to discern causal interactions ([Pearl, 2000](#)) provide a
355 way forward. Measures of functional effectiveness can, in principle, be assessed at and across
356 different levels of organisation of a system ([Marshall et al., 2016](#)), but application of these
357 rigorous theoretical measures to ecological systems demands the acquisition of extensive
358 time-series data.

359

360

361

362 As a specific example consider the (highly simplified) illustrative model community in Figure
363 1 (detailed description in Appendix), which depicts four trophic levels and five functions:
364 nitrogen supply (recycling, fixation, nitrification etc.), carbon supply, with the consequent
365 functions of carbon and nitrogen flow, and reproductive facilitation (e.g. pollination). Note
366 that the functions are individually quantified in their own *native* units (i.e. units of carbon
367 flow rate, nitrogen flow rate and e.g. pollination rate). Every function performed in the system
368 is contingent on every other, because each depends on the biomass and thereby the growth
369 rate of the species in each FEC providing it. This in turn depends on the species supplying
370 functions necessary for this growth. Analysis of functional dependencies (Figure 2) shows
371 that every function ultimately depends on the aggregate community-level biomass production
372 rate (B), but this in turn depends on every function. This illustrates two important principles:
373 first that the FECs form an autocatalytic set (*sensu* [Kauffman, 1986](#)) and second that due to
374 the recursive structure (everything depends on everything else) the system is in [causal closure](#)
375 ([Table 1](#)) and thereby ([Kauffman and Clayton, 2006](#)) constitutes a Kantian whole. The
376 closure referred to here is strictly causal and not material, nor thermodynamic. Causal closure
377 means that the network of causal interactions is sufficient to describe the behaviour at the
378 system level, given the necessary material and energy supply from the environment ([Hordijk](#)
379 [& Steel, 2004](#)), acknowledging that living systems are always materially and
380 thermodynamically open. This causal closure is a general property of the physiology of
381 individual organisms ([Luisi 2003](#)), many biochemical sub-systems of life ([Kauffman 1986](#))
382 and as our example illustrates, at least some ecological systems (for which it has important
383 implications). As an illustration of the mutual dependencies, Figure 3 shows the equilibrium
384 biomass (excluding nutrient providers) of the community which varies with a) the asymptote
385 of nutrient providers' functional response and b) the effectiveness of reproduction facilitation
386 for plants by animals. The community dynamics are tremendously complicated (Appendix) as
387 is typical of a system with several autocatalytic loops and the result of exogenous
388 perturbations is determined by internal relationships and correspondingly difficult to predict.
389

390 **The concept of Master Function**

391 Most broadly, a community that can be viewed as a Kantian whole is consistent with a
392 general definition of life: “A system can be said to be living if it is able to transform external
393 matter and energy into an internal process of self-maintenance and production of its own
394 components” (Luisi, 2003, p52). By this definition, all biological function ultimately amounts
395 to production of living cells and this is readily quantified by biomass. Accordingly, we can
396 regard biomass production as a ‘master function’ (Jaeger and Calkins, 2012), meaning that all
397 functions may be quantified by their contribution to it (by analogy, all processes in a factory
398 may ultimately be quantified by their contribution to the profit made by the factory owning
399 business). Alternatives to biomass have been inspired by approaches from physics and
400 computer science: energy dissipation based on thermodynamic theory, e.g. Kaila and Annala
401 (2008) and diversification based on complexity theory (Kauffman, 2000). We chose to focus
402 on biomass production because it is most readily measured in practice and also has a clear
403 biological and evolutionary basis. The idea of a master function removes the need for a
404 teleological account of function: it is the de-facto end point of the causal chains to which all
405 other biological functions belong, in the proximal as opposed to ultimate sense of Mayr’s
406 (1961) dichotomy (modified by Laland et al., 2011). Any biological process which enhances
407 the fitness of a biological system, increases the system’s potential to reproduce and by this,
408 over time, it increases the amount of biomass embodying it. Using biomass as the ‘master
409 function’ means that all functions at all scales of biological organisation can be quantified in
410 terms of their contribution to biomass production (which may be found *in vivo* by e.g.
411 species removal, or gene knockouts, or in computational simulations through sensitivity
412 analysis as illustrated in Figure 3). So whilst every kind of function can be quantified in its
413 own native units, we can also quantify it in the community context (referring to the function
414 to which it contributes at the higher ontological level), by its effect on aggregated biomass
415 production. Notice that this is true at the community level, but also at the level of cells, where
416 reproduction via completing the cell-cycle is the ‘master-function’ (as illustrated in Jaeger
417 and Calkins, 2012). In our suggested interpretation, biomass production is a universal
418 currency for biological function, which integrates ecological function with that defined at any
419 other level of biological organisation. We can quantify function (of any process) in terms of
420 the specific change in the rate of creation of cells (fitness), given the presence / absence of the
421 functional component (this is analogous to the power density proposed by Chaisson (2011) as
422 a surrogate for complexity). While individual processes of course all appear in different units,

423 the integrative framework proposed enables them all to be interpreted in the common
424 currency of biomass production, using e.g. mathematical simulations to examine the effect of
425 removing any functional component upon the production rate of the remaining system,
426 following recovery to a new equilibrium (as in [Fung et al., 2015](#)).
427

428 **Conclusion**

429 The Biblical lesson of the ‘tower of Babel’ was that division and confusion results from a
430 fragmentation of language. By establishing a common and precise meaning for the word
431 ‘function’, we believe a conceptual unity can arise among the various sub-disciplines of
432 biology. We have found that biological systems at all levels of organisation share the
433 properties of auto-catalytic sets and that this enables us to summarise the aggregate effect of
434 ecological functions via specific biomass production rate, which serves as an integrating
435 surrogate measure of their effectiveness. Specifically, to quantify biological function,
436 irrespective of whether this is genetic, cellular or ecological (behavioural, physiological,
437 population, community or process), we suggest the following steps:

- 438 1) Identify the ontological level of the system under study and the next higher ontological
439 level of which it is a component (Table 2).
- 440 2) Identify the FEC(s) to which the study system belongs (in relation to the higher ontological
441 level) - this amounts to specifying the system’s functions in terms of relationships among the
442 parts.
- 443 3) Perform knockout / elimination experiment (or simulation study) to quantify function in
444 terms of the master function for the higher ontological level (the whole of this level should be
445 included in the experiment).

446 This approach is now well established in molecular biology for investigating gene regulation
447 networks, the result of protein expression and signalling networks. It is developing in the
448 analysis of prokaryotic communities via functional analysis of ‘environmental’ gene sequence
449 data. Here we have shown that it can be extended to levels of organisation beyond the
450 individual organism and applied generally in all fields of ecological science, thus integrating
451 the concept of function and its analysis over all levels of biological organisation.

452
453
454
455
456

457 **Acknowledgements**

458 This work was derived from discussions held at the Synthesis Centre for Biodiversity
459 Sciences (sDiv), which inspired and funded the meeting. L. Albantakis received support from
460 the Templeton World Charities Foundation Grant TWCF 0067/AB41.

461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494

Table 1. Glossary of Terms

Causal effect: An effect that can unambiguously be attributed to a process: it only exists when the process is operative.

Cummins function: A process performed by a component that is necessary for performance of a higher level process by the system of which it is part.

Functional Equivalence Class (/Set) (FEC): The set of components of a system that share in common a particular function in the context of a higher ontological level.

Kantian whole: A self-contained system in the sense that the parts make the whole and the whole makes the parts (from raw materials and energy that may be obtained from its environment). In biology, this concept is most useful for characterising auto-catalytic sets in which components collectively by their functions, maintain and create the system to which they belong. The system is interpreted as a higher ontological level than the component parts. The information embodied by the system (at the higher ontological level) which is responsible for the ordering of the interactions among its components at the lower ontological level has been termed a *Transcendent Complex* (Farnsworth et al., 2016b).

Causal Closure: In this context (rather than in philosophy in general) closure is defined by Hordijk and Steel (2004) as the condition of an autocatalytic system whereby every reaction is catalysed by at least one component of the system and every reactant is created from a specified ‘food set’ of raw materials, using the reactions of the set. This last condition explicitly acknowledges that the system must be materially open. In general a system may be thermodynamically and materially open, yet causally closed. The Hordijk and Steel (2004) definition gives mathematical rigour to the concept of Kantian whole, i.e. a self-sustaining system supported by a sufficient supply of raw materials and energy: it is causally closed but not thermodynamically nor materially closed.

495
496 Table 2.
497
498

-
- 499 1) Molecule (including physical structure of informational bio-polymer).
 - 500 2) Molecular network (including information content of bio-polymer, control or signalling
 - 501 system and metabolic pathway).
 - 502 3) Prokaryotic cell and organelle of the eukaryotic cell and a population of such.
 - 503 4) Eukaryotic cell and a population* of such.
 - 504 5) Multicellular organism and a population* of such.
 - 505 6) Ecological community, including cellular colony and meta-community.
 - 506 7) The global bio-geo-chemical system of plant Earth.
-

508
509 * Meta-population is also included, unless inter-population interactions are complex, in
510 which case, the meta-population would be better categorised as a community.

511
512 Table 2: An illustrative ontological hierarchy of biological systems. At each level, properties
513 emerge that cannot be explained by reference to its component parts (existing at the next
514 lower level) alone. This is because every ontological level is associated with a causal
515 interaction network among the component parts. The information embodied in this network is
516 the source of the emergent properties.

517 FIGURE LEGENDS

518

519 Fig 1. An illustrative ecological community consisting of primary producer FECs (P_1, P_2);
520 primary consumers (C_1, C_2), secondary consumers (C_3, C_4 and C_5) and nitrogen suppliers (N).
521 Thin solid arrows show trophic flows, thin dotted arrows show reproductive facilitation (e.g.
522 pollination and seed dispersal), thick black dashed arrows show nutrient flow (in isolation)
523 and grey dashed show waste nitrogen, which is assumed all to return to N. The system can be
524 envisaged as three parallel functional networks (nutrient recycling, carbon pumping and
525 reproductive facilitation). Together, these functions catalyse one another so that the whole
526 system is an autocatalytic-set.

527

528 Fig. 2. Functional dependencies show the community to be an autocatalytic set. In (a)
529 reproductive facilitation (R) depends on FECs C_1 and C_4 , which in turn depend on biomass
530 production of P_1 and C_2 , respectively, which in turn depend on nitrogen supply from N and
531 reproductive facilitation R - forming an autocatalytic loop. The ellipse symbols indicate
532 'jointly necessary', otherwise parallel inputs are mutually degenerate, or redundant. C: carbon
533 supply; N: nitrogen supply, B(E) is biomass production rate of the whole community. Note
534 that all functions ultimately depend on - and result in - B(E). For this reason, B(E) provides a
535 suitable 'master function' with which to quantify all internal functions.

536

537 Fig. 3. How equilibrium total system biomass B - the master function - varies with the rate of
538 plant reproductive facilitation by consumer species and also the asymptote of the functional
539 response of nutrient providers, to supply of raw materials (nutrient FR). The plot was formed
540 from finite-difference numerical calculations of the system (Appendix equations 6-13). The
541 dynamics did not converge to an equilibrium for combinations of values, which is the reason
542 for the few missing and aberrant results (intentionally left visible in the plot).

543 **References**

544

545 Auletta, G., et al. 2008. Top-down causation by information control: From a philosophical
546 problem to a scientific research programme. - *J. R. Soc. Interface* 5, 1159–1172.

547

548 Bascompte, J. 2009. Disentangling the Web of Life. - *Science* 325: 416-419.

549

550 Baumgärtner, S. 2007. The insurance value of biodiversity in the provision of ecosystem
551 services. - *Nat. Resour. Model.* 20:87-127471.

552

553 Bohmann, K., A. et al. 2014. Environmental DNA for wildlife biology and biodiversity
554 monitoring. - *Trends Ecol. Evol.* 29:358-367.

555

556 Bruno, J. F. et al. 2003. Inclusion of facilitation into ecological theory. - *Trends Ecol. Evol.*
557 18, 119–125.

558

559 Cardinale, B. J., et al. 2012. Biodiversity loss and its impact on humanity. - *Nature* 486, 59–
560 67.

561

562 Chaisson, E. J. 2011. Energy rate density as a complexity metric and evolutionary driver. -
563 *Complexity* 16:27-40.

564

565 Chao, A., et al. 2014. Unifying Species Diversity, Phylogenetic Diversity, Functional
566 Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers. - *Ann.*
567 *Rev. Ecol. Evol. S.* 45:297-324.

568

569 Costanza, R. and Daly, H. E. 1987. Toward an ecological economics. - *Ecol. Model.* 38:1-7.

570

571 Cummins, R. 1975. Functional analysis. - *J. Philos.* 72(20):741–765.

572

573 De Groot, R.S., et al. 2012. Global estimates of the value of ecosystems and their services in
574 monetary units. - *Ecosyst. Serv.* 1, 50–61.

575

- 576 Deng, Y., et al. 2012. Molecular ecological network analyses. - BMC Bioinformatics 13:113.
577
- 578 Deyle, E. R. et al. 2016. Tracking and forecasting ecosystem interactions in real time. - Proc.
579 R. Soc. B-Biol. Sci. 283(1822).
580
- 581 Dick, J. T. A. et al. 2014. Advancing impact prediction and hypothesis testing in invasion
582 ecology using a comparative functional response approach. - Biol. Invasions 16:735-753.
583
- 584 Doolittle, R. F. 1994. Convergent evolution: the need to be explicit. - Trends Biochem. Sci.
585 19:15-18.
586
- 587 Doolittle W. F. et al. 2014. Distinguishing between "Function" and "Effect" in genome
588 biology. - Genome Biol. Evol. 6: 1234-1237.
589
- 590 Farnsworth, K. D. et al. 2012. Functional Complexity: The source of value in biodiversity. -
591 Ecol. Complex. 11: 46–52.
592
- 593 Farnsworth, K. D. et al. 2013. Living is Information Processing: From Molecules to Global
594 Systems. - Acta Biotheor. 62: 203-222.
595
- 596 Farnsworth, K.D. et al. 2016a. The complexity of biodiversity: A biological perspective on
597 economic valuation. - Ecol. Econ. 120, 350-354.
598
- 599 Farnsworth, K. D. et al. 2016b. Living through downward causation: from molecules to
600 ecosystems, in S.I. Walker, G.F.R. Ellis and P.C.W. Davies (eds) From Matter to Life:
601 Information and Causality. - Cambridge University Press.
602
- 603 Fierer, N. et al. 2012. Cross-biome metagenomic analyses of soil microbial communities and
604 their functional attributes. - Proc. Natl. Acad. Sci. U. S. A. 109, 21390–21395
605
- 606 Fung, T. et al. 2015. Impact of biodiversity loss on production in complex marine food webs
607 mitigated by prey-release. - Nat. Comm. 6. (no page numbers with this journal)
608

609 Gamfeldt, L. et al. 2008. Multiple functions increase the importance of biodiversity for
610 overall ecosystem functioning. - *Ecology* 89:1223-1231.
611

612 Ginsborg, H. 2006. Kant's biological teleology and its philosophical significance. In Bird, G.
613 (Ed.) *A Companion to Kant*. - Wiley-Blackwell.
614

615 Gould S. J. and Lewontin R. C. 1979. The spandrels of San Marco and the Panglossian
616 paradigm: a critique of the adaptationist programme. - *Proc R. Soc. B.* 205:581–598.
617

618 Gotelli, N. J. et al. 2012. Environmental proteomics, biodiversity statistics and food-web
619 structure. - *Trends Ecol. Evol.* 27:436-442.
620

621 Grey et al. 2014. Ecological networks: the missing links in biomonitoring science. - *J. App.*
622 *Ecol.* 51:1444–1449.
623

624 Griffiths, P. E. 1993. Functional analysis and proper functions. - *Brit. J. Philos. Sci.*, 44:409–
625 422.
626

627 Hooper, D.U. et al. 2005. Effects of biodiversity on ecosystem functioning: a consensus of
628 current knowledge. - *Ecol. Monogr.* 75, 3–35.
629

630 Hordijk, W. and Steel, M. 2004. Detecting autocatalytic, self-sustaining sets in chemical
631 reaction systems. - *J. Theor. Biol.* 227, 451-461.
632

633 Howe, A. C. et al. 2014. Tackling soil diversity with the assembly of large, complex
634 metagenomes. - *Proc. Natl. Acad. Sci. U. S. A.* 111:4904-4909.
635

636 Huttenhower et al. 2012. Structure, function and diversity of the healthy human microbiome.
637 - *Nature* 486:207-214.
638

639 Isbell, F. et al. 2011. High plant diversity is needed to maintain ecosystem services. - *Nature*
640 477:199-202.
641

- 642 Jax, K. 2005. Function and "functioning" in ecology: what does it mean? - *Oikos* 3:641-648.
643
- 644 Jaeger, L. and Calkins, E. 2012. Downward causation by information control in micro-
645 organisms. - *Interface Focus* 2, 26-41.
646
- 647 Jost, L. 2006. Entropy and diversity. - *Oikos* 113:363–75.
648
- 649 Kaila, V.R.I. & Annala, A. 2008. Natural selection for least action. - *Proc. R. Soc. A* 464,
650 3055–3070.
651
- 652 Kauffman, S. A. 1986. Autocatalytic sets of proteins. - *J. Theor. Biol.* 119:1-24.
653
- 654 Kauffman, S. A. 2000. *Investigations*. - Oxford University Press.
655
- 656 Kauffman, S. A. and Clayton, P. 2006. On emergence, agency and organisation. - *Phil. Biol.*
657 21:501-521.
658
- 659 Kefi, S., et al. 2012. More than a meal ... integrating non-feeding interactions into food webs.
660 - *Ecol. Lett.* 15:291-300.
661
- 662 Kellis M, et al. 2014. Defining functional DNA elements in the human genome. - *Proc. Natl.*
663 *Acad. Sci. U S A.* 111:6131–6138.
664
- 665 Knack, J. J., et al. 2015. Microbiomes of streptophyte algae and bryophytes suggest that a
666 functional suite of microbiota fostered plant colonization of land. - *Internat. J. Plant Sci.*
667 176:405-420.
668
- 669 Krause, S. et al. 2014. Trait-based approaches for understanding microbial biodiversity and
670 ecosystem functioning. - *Front. Microbiol.* 5. Article 251.
671
- 672 Laland, K.N., et al. 2015. The extended evolutionary synthesis: its structure, assumptions and
673 predictions. - *Proc. R. Soc. B.* 282: 2015101
674

675 Laland, K. N., et al. 2011. Cause and Effect in Biology Revisited: Is Mayr's Proximate-
676 Ultimate Dichotomy Still Useful? - *Science*. 334.6062 : 1512-1516.
677

678 Loreau M. 1996. Coexistence of multiple food chains in a heterogeneous environment:
679 interactions among community structure, ecosystem functioning, and nutrient dynamics. -
680 *Math. Biosci.* 134: 153–188.
681

682 Loreau, M. 2010. From populations to ecosystems: Theoretical foundations for a new
683 ecological synthesis. - Princeton University Press.
684

685 Lovelock, J. E. and Margulis, L. 1974. Atmospheric homeostasis by and for the biosphere:
686 The Gaia hypothesis. - *Tellus* 26:2–10.
687

688 Luisi, P. L. 2003. Autopoiesis: a review and a reappraisal. - *Naturwissenschaften* 90:49–59.
689

690 Marshall, W., et al. 2016. Black-boxing and cause-effect power. - arXiv 1608.03461.
691

692 Maclaurin, J. and Sterelny, K. 2008. *What Is Biodiversity?* - Univ. Chicago Press.
693

694 Mayr, E. 1961. Cause and effect in biology. - *Science* 134: 1501.
695

696 McGhee, G. R. 2011. *Convergent evolution: limited forms most beautiful.* - M.I.T.
697 Press.
698

699 Namba, T. 2015. Multi-faceted approaches toward unravelling complex ecological networks.
700 - *Popul. Ecol.* 57:3-19.
701

702 Neander, K. 2011. *Routledge Encyclopedia of Philosophy (Online).* - Routledge.
703

704 O’Gorman, E. J. and Emmerson, M. C. 2009. Perturbations to trophic interactions and the
705 stability of complex food webs. - *Proc. Natl. Acad. Sci. U. S. A.* 106 (32), 13393–13398.
706
707

- 708 Odling-Smee, F.J., et al. 1996. Niche construction. - *Am. Nat.* 147, 641 – 648.
709
- 710 Queiros, A. et al. 2011. Context dependence of marine ecosystem engineer invasion impacts
711 on benthic ecosystem functioning. - *Biological Invasions.* 13:1059-1075.
712
- 713 Pearl J. 2000. *Causality: models, reasoning and inference.* - Cambridge Univ. Press.
714
- 715 Rillig, M.C. et al. 2015. Interchange of entire communities: microbial community
716 coalescence. - *Trends Ecol. Evol.,* 30, 470–476.
717
- 718 Schleuning, M. et al. 2015. Predicting ecosystem functions from biodiversity and mutualistic
719 networks: an extension of trait-based concepts to plant-animal interactions. - *Ecography*
720 38:380-392.
721
- 722 Shipley, B. 2000. *Cause and Correlation in Biology.* - Cambridge Univ. Press.
723
- 724 Stouffer, D. B. et al. 2012. Evolutionary Conservation of Species' Roles in Food Webs. -
725 *Science.* 335:1489-1492.
726
- 727 Sugihara, G. et al. 2012. Detecting Causality in Complex Ecosystems. - *Science.*
728 338:496-500.
729
- 730 Thébault E., Loreau, M. 2003. Food-web constraints on biodiversity-ecosystem functioning
731 relationships. - *Proc. Natl. Acad. Sci. USA* 100:14949–14954.
732
- 733 Tringe, S. G. et al. 2005. Comparative metagenomics of microbial communities. - *Science.*
734 308:554-557.
735
- 736 Tononi G. et al. 1999. Measures of degeneracy and redundancy in biological networks. -
737 *Proc. Natl. Acad. Sci. U. S. A.* 96: 3257–3262.
738
- 739 Ulanowicz R. E. et al. 2014. Limits on ecosystem trophic complexity: insights from
740 ecological network analysis. - *Ecol. Lett.* 17: 127–136.

741

742 Valiente-Banuet, A., et al. 2015. Beyond species loss: the extinction of ecological interactions
743 in a changing world. - *Funct. Ecol.* 29:299-307.

744

745 Violle, C., et al. 2007. Let the concept of trait be functional! - *Oikos*. 116:882-92.

746

747 Warnecke, F. et al. 2007. Metagenomic and functional analysis of hindgut microbiota of a
748 wood-feeding higher termite. - *Nature*. 450:560-U17.

749

750 Yachi, S. and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating
751 environment: The insurance hypothesis. - *Proc. Natl. Acad. Sci. U. S. A.* 96:1463-1468.

752

753 Zenil, H., et al. 2016. Methods of information theory and algorithmic complexity for network
754 biology. - *Seminars in Cell & Developmental Biology* 51: 32–43.

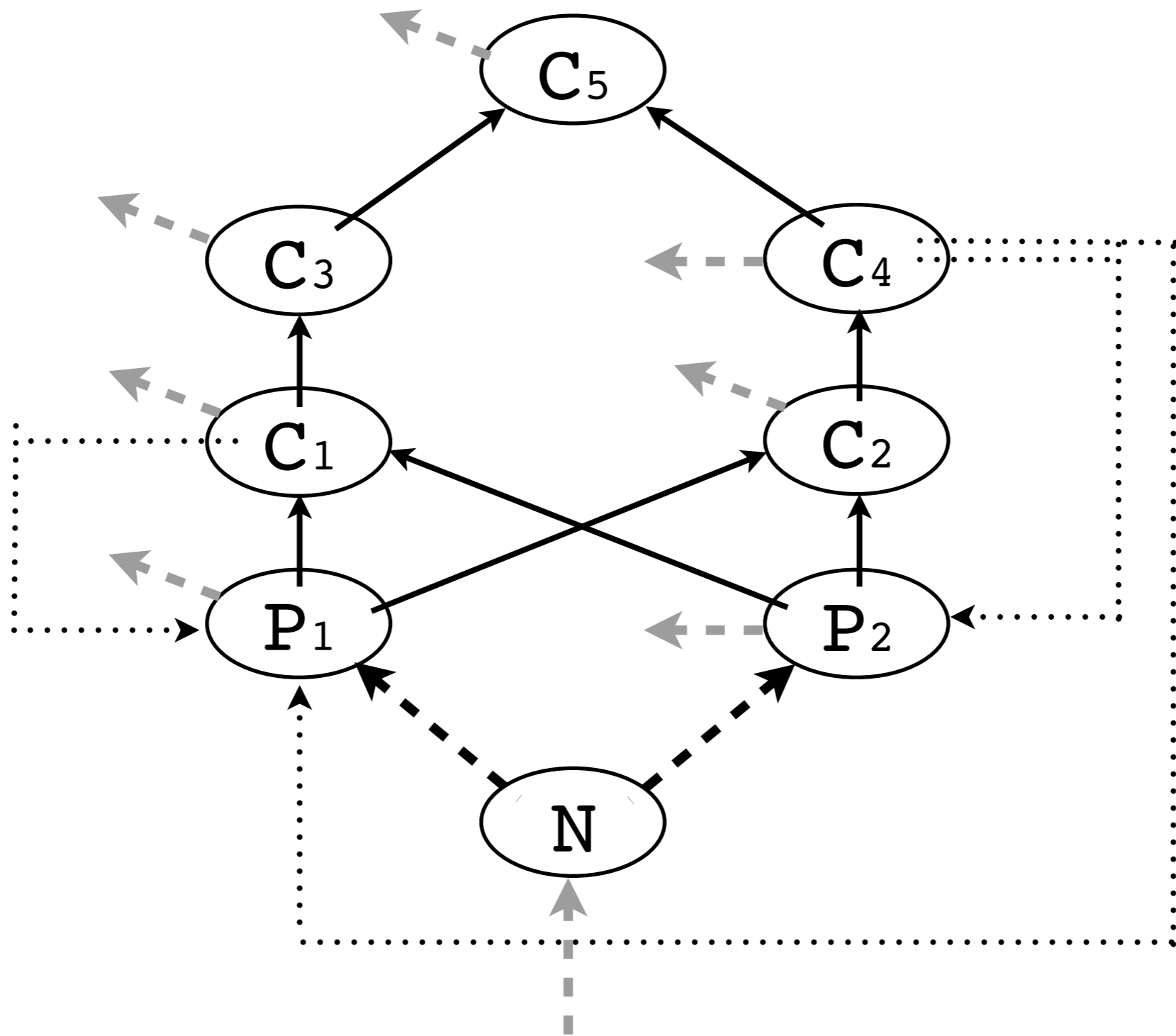
755

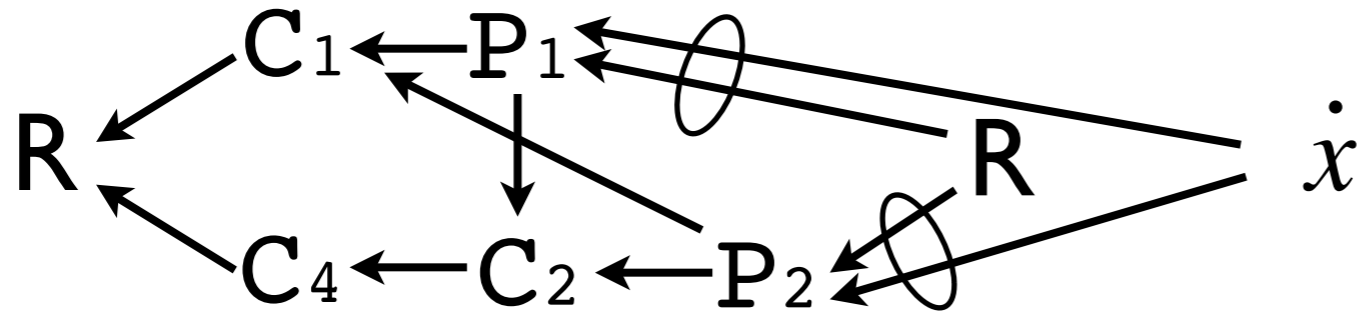
756 Zhou, J. Z. et al. 2010. Functional Molecular Ecological Networks. - *mBio* 1. (there are no
757 pages numbers with this online journal).

758

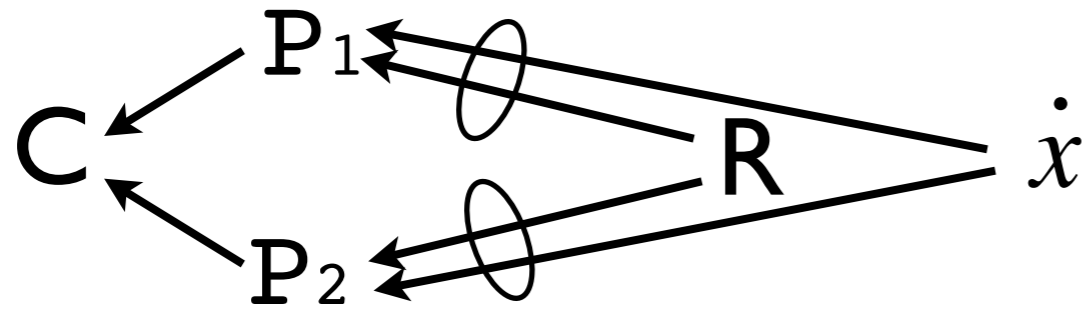
759

760

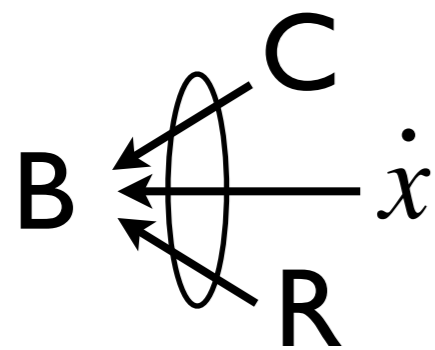




Nutrient flow \dot{x} catalyses reproduction facilitation R



\dot{x} with R catalyse carbon flow C



\dot{x} with R and C catalyse system biomass B



System biomass catalyses nutrient flow

