frontiers | Frontiers in Ecology and Evolution

TYPE Systematic Review
PUBLISHED 10 February 2026
DOI 10.3389/fevo.2026.1752597

Abundant empirical evidence of multilevel selection revealed by a bibliometric review

César Marín[1,2]*, Anne B. Clark[3], Conner S. Philson[4], Omar Tonsi Eldakar[5] and Michael J. Wade[6]

[1]Centro de Investigación e Innovación para el Cambio Climático (CiiCC), Universidad Santo Tomás, Valdivia, Chile, [2]Amsterdam Institute for Life and Environment, Section Ecology & Evolution, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [3]Department of Biological Sciences, Binghamton University, Binghamton, NY, United States, [4]Natural Reserve System, University of California, Santa Barbara, Santa Barbara, CA, United States, [5]Department of Biological Sciences, Nova Southeastern University, Fort Lauderdale, FL, United States, [6]Department of Biology, Indiana University, Bloomington, IN, United States

Natural selection is based on the concept of differential reproduction between entities, often characterized as a struggle between individual organisms. However, natural selection can act at all levels of biological organization, thus being termed "multilevel selection" (MLS). A common misconception is that selection across levels of biological organization lacks empirical support. To address this, we conducted a bibliometric review of 2,950 Web of Science/Scopus-indexed scientific articles, to document the range of taxa and research topics where MLS has been used to understand natural selection across levels. The 280 studies providing empirical support for selection at more than one level spanned a vast range of organisms, from viruses to humans to eusocial insects. They included research done both in natural populations (100) and in laboratory experiments (180). While 90.4% of studies focused on selection among organismal groups (e.g., demes, colonies, aggregates), another 9.6% explored selection across other levels (genetic elements, nuclei, cells, or multispecies communities). We classified studies by topic including artificial selection, breeding through group selection, indirect genetic effects, and contextual analysis, among others. Contrary to common notions, we found solid empirical support for the utility and importance of MLS in explaining natural selection and evolution.

KEYWORDS

artificial selection, breeding, contextual analysis, cultural evolution, indirect genetic effects, multicellularity, units of selection, group selection

EDITED BY
Sabine Nöbel,
Martin Luther University of Halle-Wittenberg, Germany

REVIEWED BY
Tomas Veloz,
Vrije University Brussels, Belgium
W. Ford Doolittle,
Dalhousie University, Canada

*CORRESPONDENCE
César Marín
✉ cmarind@santotomas.cl

RECEIVED 23 November 2025
ACCEPTED 05 January 2026
PUBLISHED 10 February 2026

CITATION
Marín C, Clark AB, Philson CS, Eldakar OT and Wade MJ (2026) Abundant empirical evidence of multilevel selection revealed by a bibliometric review.
Front. Ecol. Evol. 14:1752597.
doi: 10.3389/fevo.2026.1752597

1 Introduction

Multilevel selection (MLS) occurs when natural selection simultaneously acts at two or more levels of biological organization (Damuth and Heisler, 1988; Okasha, 2006; Wilson and Wilson, 2007; Marín, 2024). Specifically, MLS occurs when there is differential reproduction of groups in addition to reproduction of individual entities within them, or when the differential reproduction of individuals is based on their group composition or characteristics, e.g., the social environment (Goodnight et al., 1992) (see key definitions in Table 1). Goodnight et al. (1992) have defined MLS as: "variation in the fitness of individuals that is due to both properties of the individuals and properties of the group or groups of which they are members" (p. 745). Goodnight et al. (1992) definition incorporates models that explicitly include differential extinction of entire groups (Levins, 1970), trait-group models (Wilson, 1975), and Wright's (1945) definition of interdemic selection, which does not require group extinction.

The MLS framework has been useful, even essential in studying the central dogma in molecular biology (Takeuchi and Kaneko, 2019), horizontal gene transfer in bacteria (Lee et al., 2022), the origin of multicellularity (Bozdag et al., 2023), cancer (Aktipis et al., 2015), disease/virus evolution (Blackstone et al., 2020), animal (Craig and Muir, 1996) and plant breeding (Zhu et al., 2019b), as well as economics (Wilson et al., 2020) and cultural institutions (Wilson et al., 2023). The clear value of an MLS approach, whether related to the selection of particular traits, or to the discovery of what affects fitness in a given system/organism, is its focus on identifying both the direction and strength of selection from multiple sources. Despite this, criticisms and skepticism persist among biologists (Eldakar and Wilson, 2011) – albeit anthropologists seem to favor an MLS framework, according to a survey by Yaworsky et al. (2015). Marín (2024) has identified three main arguments in favor of MLS: first, the term "unit of selection" has a polysemic nature, with at least three different meanings: "interactors", "replicators or reproducers or reconstitutors", and "manifestors of accumulated adaptations" (Suárez and Lloyd, 2023; Lloyd, 2024). The second is the fact that biological entities as complex as organisms or genes must-at least-have evolved from less complex entities (Okasha, 2006). And third, there is vast empirical evidence for this theory both in laboratory and natural populations. Sound literature reviews of such empirical evidence of MLS can be found in: Wilson and Sober (1994), Goodnight and Stevens (1997), Eldakar and Wilson (2011), Goodnight (2015), Marín (2015, 2016, 2024), and Hertler et al. (2020). Despite these clear reviews and a diversity of empirical studies across a range of taxa, the misconception that MLS lacks empirical support persists (Harms et al., 2023). Here we address this misconception head on, by revealing an abundance (not a paucity) of examples of MLS in a diversity of taxa and biological systems, levels of biological organization, and type of research topics and tools.

In evolutionary biology, the evolution of altruism or prosociality has been a main focus of MLS debates for decades, but altruism is just one trait that can evolve via MLS. On the one hand, the classic example of the evolution of altruism considers groups within which selfish individuals out-compete altruists, while groups with more altruists contribute more offspring to the next generation than groups comprised of more selfish individuals

TABLE 1 Glossary of terms.

Term	Definition
Artificial selection	Human goal-driven selective breeding. Humans breed whole communities (like microbiomes) or species consortia or aggregates (like yeast aggregates) for specific desired traits (like bigger colony size, for yeasts), under imposed environmental conditions.
Breeding through group selection	Artificial selection where humans control the context for reproduction in such a way as to influence how groups of organisms function (e.g., reduced competition). Typically, these studies have two contrasting breeding treatments: individual-based breeding (classical way to breed animals or crops) and group-based breeding. "Group-based breeding" means that emergent or contextual or group-level traits are the basis for the breeding program.
Contextual analysis	Contextual analysis follows the methods for analyzing phenotypic selection originally developed by Lande and Arnold (1983) and Arnold and Wade (1984), where a multiple regression of relative fitness on phenotype is performed (Goodnight et al., 1992). Contextual analysis extends such methods by including "contextual" or "emergent" traits, that is, traits measured on the group or neighborhood, in the multiple regression. In this way, relative fitness is a function of individual and group or emergent traits. This phenotypic selection tool allows to disentangle the strength and direction of selection operating at the individual and group levels. Goodnight et al. (1992) has shown that contextual analysis is an useful tool, compatible with models that explicitly include differential extinction of entire groups (Levins, 1970), Wright's definition of interdemic selection – which does not require group extinction (Wright, 1945) – and trait-group models (Wilson, 1975).
Cultural multilevel selection	Multilevel selection (MLS) in which the inheritance system is cultural transmission, not genetic material. These studies investigated MLS in cultural traits, thus, for example these studies showed traits that confer a group-level advantage can spread via cultural MLS.
Indirect genetic effects (IGEs)	An IGE has been defined as the "effect of a gene in the genome of one individual on the phenotype of another individual" (Wade, 2026). IGEs sometimes are also deemed as "social genetic effects". A recent meta-analysis on this subject was recently published by Santostefano et al (Santostefano et al., 2025). Bijma and Wade (2008) have shown that when IGEs are included when calculating the response to selection, MLS without relatedness can explain the evolution of social traits.
Multilevel selection	MLS has been defined as a situation in which natural selection occurs among entities at two or more different levels in a nested biological hierarchy (Damuth and Heisler, 1988). Specifically, MLS occurs when there is differential reproduction of entire groups (as well as of individual entities within them), or when the differential reproduction of individuals is based on their group composition or characteristics.
Trait groups	Trait groups (Wilson, 1975) are fitness-affecting associations between two or more individuals, regardless of the duration of the association or whether actual reproduction takes place. Selection is then acting on both individuals within groups and the groups or demes themselves.

(Darwin, 1871; Wilson and Wilson, 2007). On the other hand, MLS also occurs when emergent group traits (e.g., social network structure, density, collective colony personality, among other descriptors) have significant effects on the reproductive success of a focal individual (Damuth and Heisler, 1988; Goodnight et al., 1992; Philson et al., 2025). Such effects of emergent or contextual traits have been amply demonstrated, for example in studies of epistasis (Burch et al., 2024) and indirect genetic effects (IGEs) (Linksvayer et al., 2009; Buttery et al., 2010; Bijma, 2014; Baud et al., 2021; Santostefano et al., 2025), and using techniques such as contextual analysis (Marín, 2016; Suárez and Lloyd, 2023; Lloyd, 2024; Philson et al., 2025).

We conducted a bibliometric review of the scientific literature to identify the breadth and depth of empirical evidence of MLS across levels of biological organization. In addition, we also focused on phenotypic selection studies that use contextual analysis (Heisler and Damuth, 1987) to decompose the strength and direction of selection at different levels (individual organisms and groups of organisms). We then organized the literature on the basis of study systems (i.e., eusocial insects, humans, wild plants and algae, crops, etc.), levels of biological organization assessed (demes, communities, cells, etc.), and type of research (i.e. *in situ* studies of natural populations, artificial selection experiments, breeding through group selection, etc.). The focus of this review is to provide an introduction, accounting, and organization of the vast empirical support of MLS and its utility to understand the natural world. In this review, "support" means only that selection at different levels was explicitly measured, not that higher levels or "group" selection outweighed lower levels. For example, there were some studies in which lower-level selection or individual properties were shown to be more important than higher-level selection or properties in explaining focal individual fitness (Tsuji, 1995; Donohue, 2003; Fisher et al., 2017; Fisher et al., 2021). Thus, our bibliometric review captured the full spectrum of studies, whether or not individual or higher-level selection is the prevailing force – something perfectly consistent with MLS theory, as MLS should be evaluated in a case-by-case basis (Wilson and Wilson, 2007). In addition, while discussions of alternative and complementary frameworks (such as inclusive fitness theory or direct reciprocity) and mechanisms that partition variation within and between groups (e.g., conditional dispersal, kinship, and kin groups) are of general interest (for detailed discussions, see Bijma and Wade, 2008; Goodnight, 2013; Frank, 2025; Marín and Wade, 2025), the consideration of such topics are beyond the scope of this review. However, in Appendix 1 we discuss why MLS is not mathematically equivalent to other frameworks such as inclusive fitness theory.

2 Materials and methods

2.1 Bibliometric analysis: search terms

The current review is classified as a "bibliometric" analysis and not as a "meta-analysis" because, with the exception of the regression coefficients of 21 studies involving contextual analysis, no actual data was extracted from the articles. Rather, this review aimed at compiling the empirical evidence for MLS *in situ* and in laboratory experiments by conducting a bibliometric analysis following the "Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO)" (Montazeri et al., 2023). Please find in Appendix 2 the BIBLIO complete check-list required in such preliminary guideline.

In January 2025, the following terms were searched in the Scopus (https://www.scopus.com/home.uri) database: "multilevel selection" across the whole article, and "group selection" in the Title, Abstract, and Keywords – because the latter was the term most commonly used before Damuth and Heisler (1988). The search spanned 1900–2024 and included articles and reviews only published in English, in journals indexed both in Web of Science and Scopus. In Scopus, the following areas were excluded from the search: dentistry; nursing; energy; chemical engineering; health professions; pharmacology, toxicology and pharmaceutics; business, management, and accounting; materials science; physics and astronomy; engineering; computer science; arts and humanities; mathematics; and medicine. All the remaining areas were included in the search. We also conducted an additional search in Google Scholar, with the same terms as in the Scopus search, to capture Web of Science/Scopus-indexed MLS empirical papers not discovered by the Scopus search due to differences in both search engines.

2.2 Bibliometric analysis: identification, screening, eligibility, and inclusion criteria

The bibliometric analysis had a total of four phases: identification, screening, eligibility, and inclusion (Figure 1). In the identification phase, all duplicates were deleted, and in the screening phase, based on information contained within the abstracts, all articles not related to biology, cooperation, and social behavior in general, were excluded. For the eligibility phase, all non-empirical studies were excluded, again based on the content within the abstracts. These non-empirical studies included mathematical models, reviews, discussions, response articles, conceptual models, and opinion articles, among others. In the inclusion phase, the articles were read in their totality, and those articles indicating MLS or group selection as "possible" or "plausible" (but not surely) mechanism explaining the observed results or patterns, were also excluded. For example, among the articles excluded on this third phase is an article entitled: "Sex-ratio bias and possible group selection in the social spider *Anelosimus eximius*" published in *The American Naturalist* (Aviles, 1986), because the author indicates that group selection *might* be the mechanism explaining her results but further research is needed. All articles employing the same type of argumentation or reasoning were also excluded.

2.3 Bibliometric analysis: classification

After the inclusion phase, articles were classified according to the type of study (*in situ* or laboratory experiments); taxon or

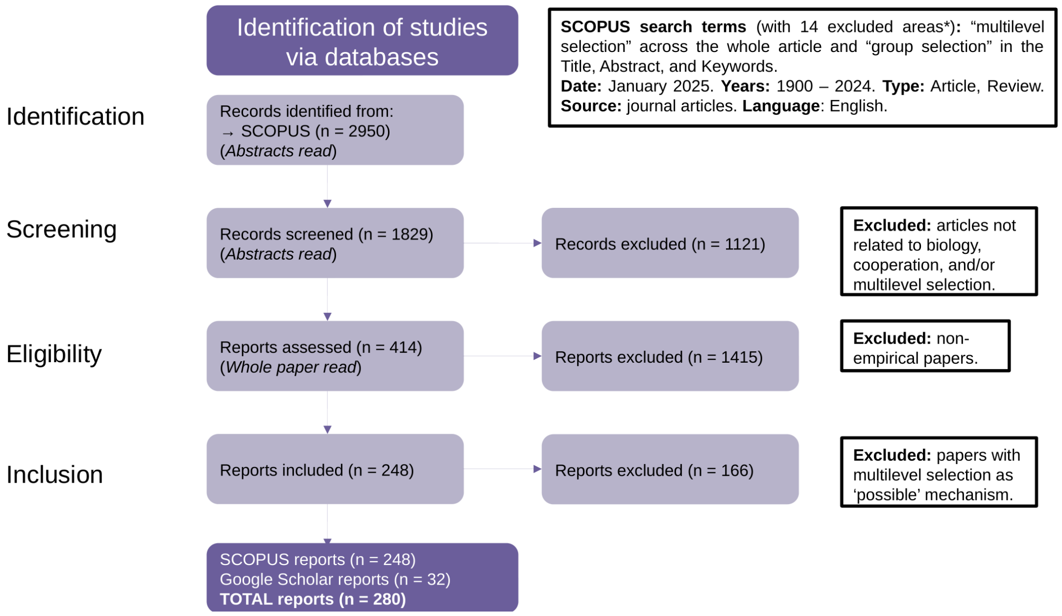

FIGURE 1
BIBLIO flow diagram for bibliometric review of empirical studies of multilevel selection.

study system (fungi, farm animals, eusocial insects, humans, microbiomes, etc.); the level of biological organization that was the main focus of research (groups or demes of organisms, communities, colonies, nuclei, aggregates, selfish genetic elements, etc.); and the main topic (or sub-topic) or method to assess MLS *in situ* or in the lab. For the latter, we identified a total of 9 topic categories and 70 sub-categories of MLS empirical research (Supplementary Table S1 in Appendix 2). A general overview and specific details, as well as information about the exclusion/inclusion criteria of each category and sub-category, can be found in Appendix 2. The full list of MLS empirical articles, after the inclusion (third) phase, can be found in Appendix 3.

The MLS *in situ* studies included five categories (a more detailed description can be found in Appendix 2), as follows:

1. *Cultural multilevel selection*: those that investigated MLS in the spread of cultural traits, and, for example, demonstrated that traits conferring a group-level advantage can spread via cultural group selection.
2. *Dataset analyses*: these studies analyzed historical or published data to infer MLS processes occurring in natural populations or communities. A subset of these studies implemented different sorts of molecular sequencing to natural populations, using different tools, from single-nucleotide polymorphism analysis to genome-wide association studies. Another subset implemented phylogenetic analyses either to assess selection at the species level or to explain the evolution of complexity/multicellularity across phylogenetic trees.
3. *Indirect Genetic Effects (IGE)*: an IGE has been defined as the "effect of a gene in the genome of one individual on the phenotype of another individual" (Wade, 2026). IGE studies collect population and trait and/or loci data to assess the effects of interacting partners on a focal individual traits' and/or reproduction.
4. *Group effects*: these studies assessed the effects of group emergent properties (like networks of interactions or group structure) on focal individuals' trait variation and/or individual fitness. A subset of these studies assessed group heritability, which has been defined as the "tendency of offspring groups to resemble their parental groups with respect to group-level traits" (Okasha, 2003). Another subset detected colony-level selection, by directly measuring phenotypic variation at the whole-colony level, in eusocial insects. Another subset still assessed group effects on focal individuals' phenotypic variation and/or fitness under field conditions.
5. *Contextual analysis*: contextual analysis extends the commonly used methods to measure natural selection in natural populations (Lande and Arnold, 1983; Arnold and Wade, 1984) by including "contextual" or "emergent" traits, that is, traits measured on the group or neighborhood, in a multiple regression. In this way, relative fitness is a function of individual and group or emergent traits.

The MLS experimental studies included four categories (a more detailed description can be found in Appendix 2), as follows:

1. *Lab experiments*: some lab experiments imposed group and individual selection regimes and compared responses to selection afterwards, some measured the molecular consequences of such treatments, others measured group effects on focal individuals' fitness, microbial culture treatments, and measurements of different aspects of

colony-level selection (trait variation, fitness, among others). A subset of these studies consisted of controlled experiments done to assess how IGEs affect focal individuals phenotypic variation and/or fitness.

2. *Breeding through group selection*: typically, these studies have two contrasting breeding treatments: individual-based breeding (classical way to breed animals or crops) and group-based breeding, measuring the individual and group phenotypic effects and productivity of both treatments after several generations. A subset of these studies were breeding programs that incorporated the calculation and effects of IGEs.
3. *Psychology experiment*: these were psychological experiments following and aimed to assess a cultural multilevel selection framework (Wilson et al., 2023).
4. *Artificial selection*: in these studies, researchers selected whole communities (like microbiomes) or species consortia or aggregates (like yeast aggregates) for specific desired traits (like bigger colony size for yeasts), under specific environmental conditions. For example, studies implementing artificial selection for multicellularity in yeasts (Ratcliff et al., 2012; Bozdag et al., 2023) match this category.

These nine categories were created by organizing all qualifying MLS empirical papers by similarity and/or main topic and/or main method assessed. The 70 sub-categories are mostly related to specific taxa or study systems, techniques, or sub-topic (Appendix 2).

2.4 Contextual analysis studies

Lastly, with the specific goal of comparing the strength and direction of natural selection as measured across different levels of biological organization, we conducted a detailed analysis of the 25 phenotypic selection studies that explicitly measured selection at multiple levels of biological organization (individual organisms and demes). A recently published article implementing contextual analysis in wild mammal populations was added to this analysis (Philson et al., 2025), totaling 26 studies. Specifically, we extracted the available beta regression coefficients of each study, as these coefficients depict the direction and strength of selection on the trait in question at individual and group levels. The complete dataset of these coefficients is found in the Supplementary Table S3 of Appendix 2.

3 Results

The identification phase of the Scopus search yielded a total of 2,950 articles (after deleting duplicates) (Figure 1). A total of 1,829 papers remained after exclusion of all articles not related to biology, cooperation, and social behavior in general (screening phase). From these, only 414 papers included empirical studies and thus persisted in the eligibility phase (Figure 1). Finally, 166 articles indicating group selection or MLS as possible or plausible mechanism but not ensuring it as an explanation, were also excluded, resulting in a total of 248 papers providing empirical support for MLS found with Scopus. The additional search with Google Scholar, which was restricted to Web of Science and Scopus-indexed articles, added 32 articles to this list, leading to a total of 280 scientific articles providing empirical support for MLS (Figure 1).

These articles spanned 1976–2024, and 180 consisted of laboratory-controlled experiments, while the remaining 100 consisted of *in situ* (field) measurements and/or experiments (Figure 2). Only years 2019, 2021, and 2023, yielded 20 or more MLS empirical papers, with a peak of 22 studies in 2019 (Figure 2). Only 81 studies were published during the first 35 years of MLS empirical research (1976–2011), while the remaining 199 have been published since 2012, showing a marked increase in research in the last 12 years, both on MLS *in situ* and experimental studies (Figure 2).

Regarding the taxa or study systems, systems like farm animals, eusocial insects, “other insects” (this means non-eusocial insects such as beetles, spiders, water striders, among others), and humans, together compose approximately 65% and 55% of experimental and MLS *in situ* studies, respectively (Figure 3). However, in MLS *in situ* studies, systems like wild plants and algae, wild mammals, and wild birds also make up an important proportion of studies, while this is the case for microbial populations and fungi in MLS experimental studies (Figure 3). Many other study systems or taxa have been empirically investigated under a MLS framework: other invertebrates (such as tunicates and polychaetes), crops, algae, fish and amphibians, microbiomes, etc. (Figure 3).

Regarding the levels of biological organization investigated, 90.4% of MLS empirical studies focused on individual organisms and groups of organisms. In particular, 71% of studies (198 papers) focused on demes, while 19.4% of studies focused on tighter organismal groups: 24 studies were conducted at the “aggregate” level (aggregates of bacteria, amoebas, algae, and yeast) and 31 studies investigated colony-level selection (mostly in eusocial insects but also including spider and *Caenorhabditis elegans* colony studies). A 9.6% of MLS empirical studies focused on organization levels above or below organisms/groups of organisms: four studies were conducted at the cell level (this include horizontal gene transfer or RNA viruses, for example); three studies were conducted at the genetic element level (specifically investigating selfish genetic elements or gene transfer agents using an MLS framework); 13 studies investigated community-level selection (mostly microbiomes, but also including beetles, ants, and arthropod communities); three studies with either algae or sea-grass investigated clonal or module-level selection, i.e. selection acting at the clonal level; two studies with fungi investigated natural selection at the nuclei level, as some fungal taxa can contain thousands of nuclei on a single spore (Jany and Pawlowska, 2010); and finally, two studies investigated natural selection at the species level.

Both *in situ* and experimental MLS empirical evidence comes from many different sources, types of study, and taxa or study systems, to the point that our nine main categories were sub-categorized into 70 sub-categories (Figure 3; Appendix 2). More than half (n=54) of MLS *in situ* studies used either IGEs

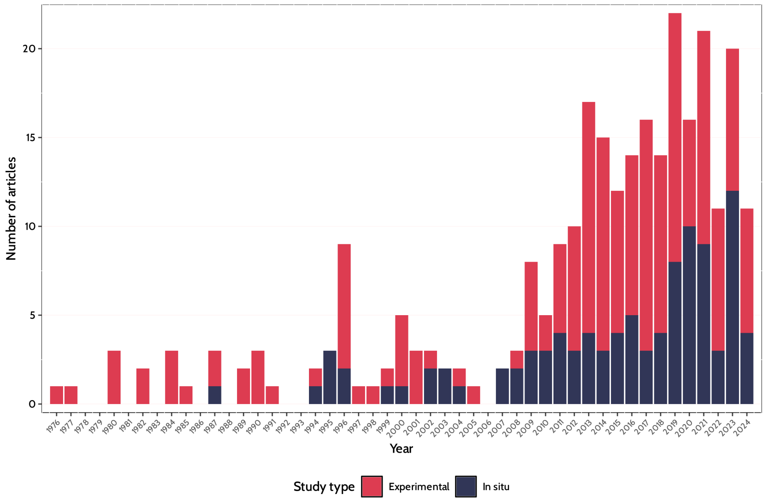

FIGURE 2
Number of Web of Science/Scopus-indexed articles (n=280) published between 1976 and 2024 providing empirical support for multilevel selection *in situ* (n=100; blue) and through Experimental studies (n=180; red).

measurements or contextual analysis, with categories such as group effects (n=18) and dataset analyses (n=18) also having important numbers (Figure 3). Similarly, 107 of the 180 MLS experimental studies were laboratory experiments of different types, with the group selection treatments on wild animals sub-category in particular having 18 studies (Appendix 2). Other MLS experimental studies categories also had an important number of articles, including breeding through group selection (n=32), artificial selection (n=22), and psychology experiment (n=19) (Figure 3). A brief summary of nine representative studies of each one of the main categories is given in Table 2 for MLS *in situ* studies and in Table 3 for MLS experimental studies.

Finally, regarding the 26 studies that implemented contextual analysis in natural populations, it was not possible to extract the regression coefficient information from five of them (Appendix 2; Appendix 3). Thus, Figure 4 shows the regression coefficients from 21 studies spanning 1995–2025, which were conducted in a plethora of study systems: from plants and water striders to chipmunks and humans. In Figure 4, the effects of individual ("size") and group (average "size" of the neighborhood individuals) traits on focal individuals' fitness is shown with the Beta (β) regression coefficients. In some studies (Stevens et al., 1995), group selection is stronger and goes in an opposite direction than individual selection, while in other studies (Donohue, 2004) the strength and direction of individual and group selection are similar, and in other studies (Bolstad et al., 2012), individual selection is significantly stronger than group selection (Figure 4). In summary, there is a variety of selection outcomes across the 21 studies as revealed by contextual analysis, with some showing selection at different levels acting in concert while others show selection acting in opposition (Figure 4).

4 Discussion

In this review, we catalogued a vast body of empirical evidence for multilevel selection (MLS), from *in situ* observations to experimental studies, spanning five decades. Such evidence encompasses a broad spectrum of study systems and taxa, albeit systems like farm animals, eusocial and non-eusocial insects, and humans have been the main focus of MLS research. Similarly, and likely due to the organismal focus of most biologists, but also due to methodological feasibility, individual organisms and groups of organisms (demes, colonies, aggregates) have been the most investigated levels of selection in the MLS empirical research literature. With our analysis we can conclude that there is not a single or majority way to investigate MLS *in situ* or experimentally. Rather, multiple tools or ways of empirically investigating MLS have been used through the decades, which respond to the specificity of each study system or taxa, level of organization, topic, and/or tool. Further, our bibliometric screening shows that from 1,829 articles that deal in some way with MLS or social evolution, 1,415 articles (77%) consisted of mathematical and conceptual models (Figure 1), opinion pieces, debates, reviews, simulations, and so on. These are important in their own right but are excluded here because we are concerned with the realized utility of the MLS framework in empirical research. Group selection was initially rejected, not due to lack of evidence, but due to its supposed theoretically implausibility (Maynard Smith, 1964; Williams, 1966). The large number of models demonstrating the theoretical plausibility of MLS therefore complements our review of the empirical literature.

The debate on the units of selection has gone on for too long. It is high time to move on and focus on the empirical evidence and

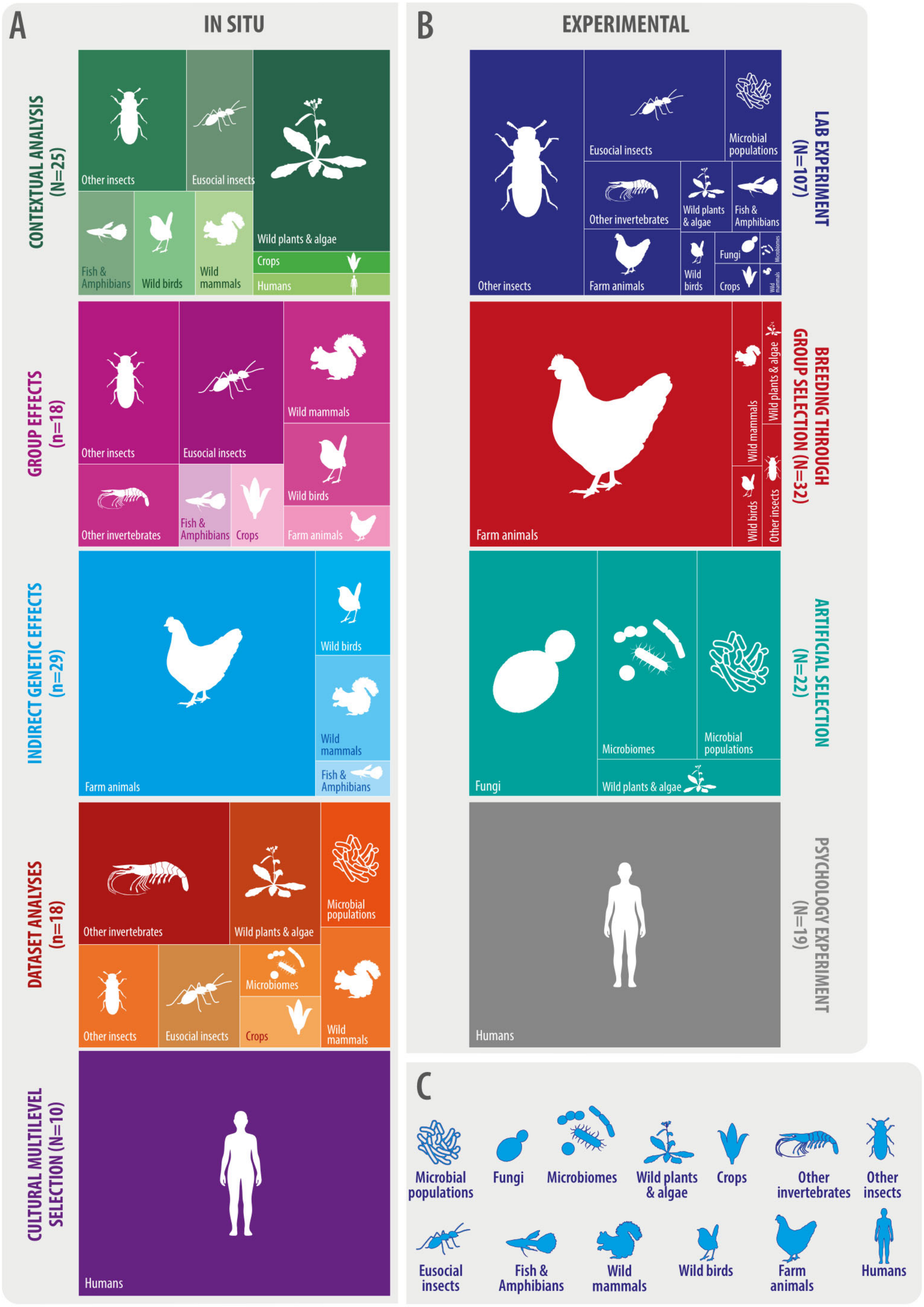

FIGURE 3
Distribution of 280 studies providing empirical support for multilevel selection across study systems and types. **(A)** *In situ* studies. **(B)** Experimental studies. **(C)** Study systems.

data (Marín, 2015; Marín, 2016; Marín, 2024). Because there is a plurality of levels of selection investigated, from selfish genetic elements and nuclei to microbiomes, not all tools or experiments will work the same. For example, it would be quite challenging to apply contextual analysis in artificial selection experiments dealing with multicellularity evolution. Following our comprehensive MLS definition (Table 1), here we focused on compiling an extensive list of studies showing either differential reproduction of entire groups, or the differential reproduction of individuals being affected by their group composition or characteristics.

The plurality of study types and systems involves a variety of different methods to assess MLS *in situ*. For example, some *in situ* studies infer MLS by using molecular sequencing tools such as microsatellites, fingerprinting, and genome-wide association studies, among other tools. More than half of MLS *in situ* studies implemented either IGE assessment or contextual analysis

TABLE 2 Representative studies of multilevel selection (MLS) *in situ* studies.

Study type and reference	Description
1. Cultural multilevel selection (Turchin and Gavrilets, 2021)	Using a database of past societies history (*Seshat: Global History Databank*), the authors found that the tempo (rates of change) of cultural macroevolution is characterized by periods of apparent stasis interspersed by rapid change. They found that the most important macroevolutionary patterns include competition and warfare but also cultural exchange and selective imitation, fully in accordance with cultural multilevel selection theory.
2. Dataset analyses (Herron and Michod, 2008)	This study investigated the transition from unicellular to multicellular life in Volvocine algae. Phylogenetic reconstructions of ancestral character states were derived from the diverse array of extant species in the volvocine lineage ranging from unicellular to colonial forms that themselves vary in size, structure, and degree of cellular specialization. Herron and Michod (2008) describe an evolutionary history with multiple independent origins and reversals of traits that underlie cellular cooperation (i.e., transition of fitness from individual cells to the group level) as well as conflict-mediation mechanisms to curtail the exploitation of cooperation.
3. Indirect Genetic Effects (IGE) (Santostefano et al., 2021)	The authors assessed how IGEs contributed to genetic variation of behavioral, morphological, and life-history traits in a wild Eastern chipmunk population, comparing the contribution of direct and indirect genetic effects to trait evolvability. They found significant IGEs for trappability and relative fecundity, but little direct genetic effects in all traits measured.
4. Group effects (Robinson et al., 2023)	The ant *Rhytidoponera metallica* forms queen-less colonies, with such a low intra-colony relatedness that they are proposed as a transient, unstable form of eusociality. Despite this, these ants are among the most widespread in Australia, showing that relatedness is not necessary for such success. The authors show that these ants exhibits remarkable intra-colony variation regarding their polypeptidic venom composition (revealed by transcriptomic and mass spectrometry), with workers sharing only a relatively small proportion of toxins in their venoms. Such variation is not due to the presence of chemical castes, but is rather explained by toxin allelic diversity. The authors conclude that such high toxin diversity is explained through MLS, selecting for colonies that can exploit more resources and defend against a wider range of predators.
5. Contextual analysis (Stevens et al., 1995)	This constitutes the first study to implement contextual analysis (Heisler and Damuth, 1987) in natural populations. This study partitioned selection into group and individual level components in natural populations of *Impatiens capensis*, measuring the relationships between three fitness components and several group and individual level traits. Two of the fitness components (survival rate and cleistogamous seed production) were affected by individual and group selection, while chasmogamous seed production (the third fitness component) was only affected by individual selection.

TABLE 3 Representative studies of multilevel selection (MLS) experimental studies.

Study type and reference	Description
1. Lab experiment (Wade, 1976)	First empirical study of MLS in our bibliometric search. Wade (1976) imposed group selection for both increased and decreased adult population size in laboratory populations of the beetle *Tribolium castaneum*, at 37-day intervals. Individual selection control treatments (i.e. no group selection imposed) were included. Response to the group selection treatments occurred fast, at three or four generations, and in general was large in magnitude (sometimes 200% larger magnitude than the control).
2. Psychology experiment (Francois et al., 2018)	This study provides evidence both from survey data and laboratory treatments of experimental subjects, consistent with a set of core concepts and theories based on cultural MLS. Specifically, the authors find that "increases in competition increase trust levels of individuals who (i) work in firms facing more competition, (ii) live in states where competition increases, (iii) move to more competitive industries, and (iv) are placed into groups facing higher competition in a laboratory experiment". They conclude that their findings provide support for cultural MLS as a contributor to human prosociality.
3. Breeding through group selection (Craig and Muir, 1996)	An important behavioral problem with egg laying hens is their proclivity to aggressively peck their cage-mates. This can be minimized through the practice of beak-trimming; however, this can cause lasting pain for the animals involved, thus essentially improving one scenario of animal well-being at the cost of another. Craig and Muir (1996) investigated whether beneficial behaviors could be selected for at the group-level, thereby eliminating the need for beak-trimming. Three genetic stocks of hens were compared for mortality, injuries, and body condition: one of the lines involved the seventh-generation of group-selected hens (recurrent selection of the most productive cages), an unselected stock of hens, and a highly productive, typically beak-trimmed commercial stock. Overall, the group-selected lineage showed behavioral improvements over the unselected and commercial lines resulting in reduced cannibalism, better feathering, and improved welfare. Furthermore, when comparing the previous six generations of the group-selected line of collectively house hens to those housed individually (Craig and Muir, 1996), by the sixth-generation the collectively housed hens approximated the mortality of their solitary counterparts (8.8% to 9.1%, respectively). This was the result of a dramatic decrease in mortality from 68% in the second generation down to 8.8% in the sixth-generation of group-selected hens. In addition, the group-selected lineage also experienced substantial improvements in survival (from 169 to 348 days) and egg production per hen (from 91 to 237 eggs) over that same time frame.
4. Artificial selection (Bozdag et al., 2023)	This multicellularity long-term evolution experiment was carried out with snowflake yeast (*Saccharomyces cerevisiae*), by selecting for larger group size under three metabolic treatments: anaerobic, obligately aerobic, and mixotrophic yeast. After 600 rounds of selection, yeast in the anaerobic treatment group evolved to be macroscopic, becoming around 2×10^4 times larger (about 1 mm, visible to the naked eye) and about 10^4-fold more biophysically tough, while retaining a clonal multicellular life cycle. Yeast in the aerobic treatment remained microscopic (only sixfold larger). This was explained through biophysical adaptation of increasingly elongate cells, which after some time facilitated branch entanglements that enabled groups of cells to stay together.

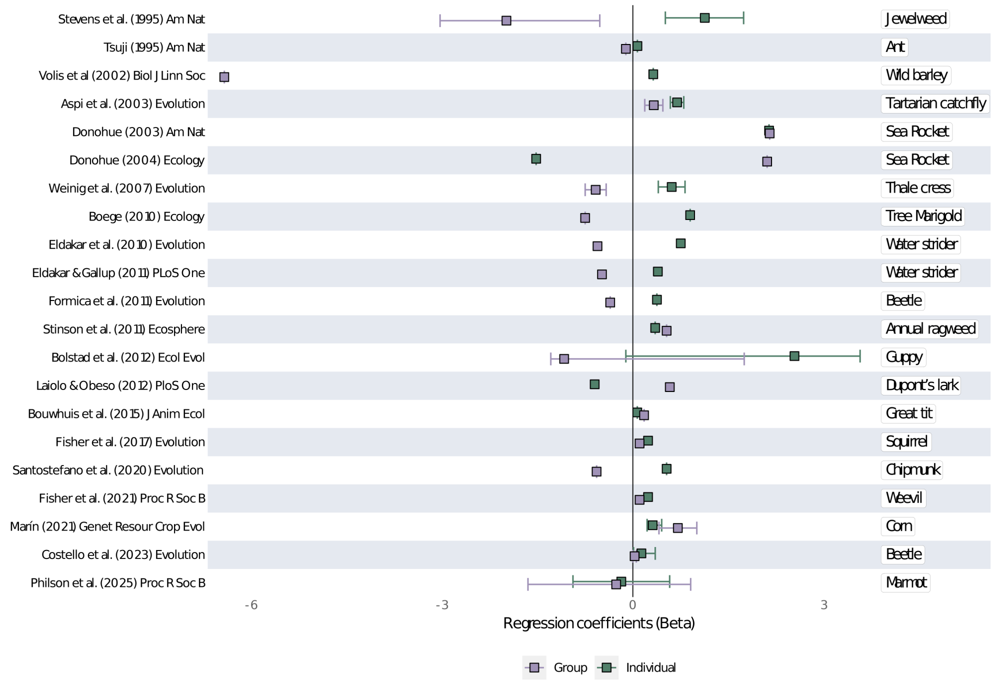

FIGURE 4
Summary of 21 (out of 26) contextual analysis of phenotypic selection done between 1995 and 2025. Beta regression coefficients show the effects of organismal "size" (or traits which are a proxy of size, like height) at the individual (green squares) and group (purple squares) levels, on individual fitness (or fitness proxies). References cited: Stevens et al. (1995), Tsuji (1995), Volis et al. (2002), Aspi et al. (2003), Donohue (2003, 2004), Weinig et al. (2007), Boege (2010), Eldakar et al. (2010), Eldakar and Gallup (2011), Formica et al. (2011), Stinson et al. (2011), Bolstad et al. (2012), Laiolo and Obeso (2012), Bouwhuis et al. (2015), Fisher et al. (2017), Santostefano et al. (2020), Fisher et al. (2021), Marín (2021), Costello et al. (2023), and Philson et al. (2025).

(Figure 3), finding quite significant effects of the neighborhood traits or emergent traits on focal individuals' fitness (and individual trait variation). Such neighborhood/emergent effects are a core feature of MLS (Table 1), with IGEs, contextual analysis, and group effects measurements, representing different ways in which they are calculated. It is not within the scope of this article to compare such mechanisms of assessment, as this has been done plenty in the literature (Bijma and Wade, 2008; Goodnight, 2013). In particular, Bijma and Wade (2008) have shown the relationships between kin selection (inclusive fitness theory), MLS, and IGEs. In Appendix 1 you can find an expanded Discussion on these relationships and on the non-equivalence between MLS and inclusive fitness theory. Rather, here we show that when group composition or characteristics or average/emergent traits are considered in the response to selection models' (in addition to individual traits), focal individuals' fitness are affected by such group composition or characteristics. This is supported by recent meta-analyses by Santostefano et al. (2025) and Burch et al. (2024), which respectively showed that IGE and epistasis are ubiquitous across the Tree of Life.

The comprehensive definition of MLS that we employ here (Table 1) falls into the disambiguating project of the units of selection literature (Suárez and Lloyd, 2023). In the disambiguating project (Suárez and Lloyd, 2023; Lloyd, 2024), the "units of selection" at any given level in the biological hierarchy can have one or more of three functional roles in the process of natural selection, which must be distinguished from each other. These roles are: interactors (phenotypic variation and differential proliferation), replicators (or reproducers or reconstitutors; inheritance), and manifestation of accumulated adaptations (Suárez and Lloyd, 2023). Thus, for a biological entity to be considered an interactor, two minimal things are required: phenotypic variation and differential proliferation. Furthermore, Suárez and Lloyd (2023, p. 17) have defined natural selection as a "process in which the differential proliferation of interactors causes the differential replication of replicators" [or the differential reproduction of reproducers or the differential reconstitution of reconstitutors]. This clarification is necessary, as many of the historical (Williams, 1966) and current-day (Harms et al., 2023) critiques of MLS confound the roles of the different units of selection (Gould and Lloyd, 1999), requiring phenotypic variation, differential proliferation, and inheritance at the same level of biological organization to be considered as a unit of selection. This is not necessarily the case. For example, although typically genes constitute replicators, in specific cases such as selfish genetic elements, genes *might* also be considered as interactors (Gitschlag et al., 2020).

The comprehensive definition of MLS (Table 1) employed here captures instances in which entire groups constitute the inheritance unit (replicator/reproducer/reconstitutor) and instances in which entire groups constitute the interactor but inheritance occurs at a lower level of biological organization (most typically, the individual organism or its genetic material). The latter cases are typically detectable with techniques such as IGE measurements, social network analysis, the Price equation, and contextual analysis, among others (Marín and Wade, 2025), as mentioned above. In summary, MLS occurs when natural selection operates simultaneously among two or more different levels of a nested biological hierarchy, which either causes differential reproduction of entire groups (i.e., the group is also the replicator/reproducer/reconstitutor) or when the differential reproduction of individuals is influenced by their group composition or its characteristics (i.e., lower-level entities are the replicator/reproducer/reconstitutor).

The pioneering study by Wade (1976) (Table 3) was the starting point of laboratory studies on which group selection was imposed as a treatment. Several dozen similar studies (imposed group selection in laboratory populations) were conducted through the decades, generally finding rapid responses to the group selection treatments after a few generations. Further, such imposed group selection studies found that selection sometimes acts in concert and sometimes in opposition at the individual and group levels, also with varying strength. Interestingly, the same pattern is found when analyzing contextual analysis studies (Figure 4): natural selection sometimes acts at the same and sometimes at different directions and strengths across levels of biological organization. As such, no generalization can be made about MLS and it should be investigated on a case by case manner (Wilson and Wilson, 2007; Eldakar and Wilson, 2011). However, ecological constraints can help predict responses to selection. For example, when in 2017 the category 4 Hurricane Maria almost totally destroyed a Puerto Rican island inhabited by rhesus macaques, shade became a very scarce resource. As a response, there was a marked increase in tolerance and decrease in aggression among macaques (Testard et al., 2024), with the most tolerant animals having the highest survival. Similarly, in plant-mycorrhizal associations it has long been known that under scarcity of nutrients (particularly nitrogen and phosphorous), this symbiotic association becomes more mutualistic while under “luxury” conditions (excess of nutrients), the usually benign mycorrhizal fungal microbiomes can behave as nutritional parasites (Johnson et al., 1997; Johnson and Marín, 2025).

Several other influential MLS experimental studies include Craig and Muir (1996), Swenson et al. (2000), Ratcliff et al. (2012), and Bozdag et al. (2023) (Table 3). Craig and Muir (1996) and several dozen more studies (a total of 32 studies; Figure 3) have shown that MLS is a very useful framework for breeding programs of farm animals and crops. Furthermore, when farm animals or crops are bred through group selection treatments (i.e., selecting group traits) or when IGEs are considered in breeding programs, the outcome is almost always the desired for the farmer: higher yields or more production. Even MLS sceptics recognize the value of MLS-focused breeding programs in wheat cultivars (Zhu et al., 2019b; Zhu et al., 2019a; Zhu et al., 2022). Empirical evidence showing the success of wheat breeding for higher yields over the past 100 years in north-western China has been argued to result in part from “unconscious group selection on root traits” (Zhu et al., 2019b), which results in smaller, less branched, and deeper roots.

Swenson et al. (2000) pioneered the framework of artificial ecosystem selection as a way of selecting communities of soil microorganisms based on plant performance. This implies exposing multiple generations of plants to particular selection pressures, selecting the microbiomes that increase plant fitness (or selected traits) to the next generation, while the genetic basis of the host remains the same. This approach has been successfully used to engineer belowground communities that increase plant tolerance to drought (Lau and Lennon, 2012; Jochum et al., 2019) and salinity (Mueller et al., 2021), or that increase leaf greenness (Jacquiod et al., 2022), among others (reviewed in Sánchez et al., 2021; Sanchez et al., 2023; Yu et al., 2023). On their part, Ratcliff et al. (2012) and Bozdag et al. (2023), implementing artificial selection regimes in yeast aggregates, have shown some of the most visually stunning examples of experimental MLS: they shown *de novo* evolution of macroscopic multicellularity just after one year and 600 rounds of selection (Bozdag et al., 2023). In particular, in an anaerobic treatment, yeast evolved to be macroscopic, becoming 2 x 10^4 times larger than at the beginning, while maintaining a clonal multicellular life cycle (Bozdag et al., 2023).

A MLS framework has long been used to investigate human culture (Soltis et al., 1995), originating a whole sub-discipline, deemed “cultural multilevel selection” (Wilson et al., 2020; Wilson et al., 2023). In our review, a total of 30 MLS empirical studies were centered around humans: 19 consisted of psychological experiments, 10 assessed or inferred cultural MLS *in situ*, and one implemented contextual analysis over 55 years of polygyny and polyandry data, based on the Utah Population Database (Moorad, 2013). MLS seems to explain the most important cultural macro-evolutionary patterns and historical trends, including competition and warfare but also exchange and selective imitation (Turchin and Gavrilets, 2021) (Table 2). The utility of MLS has been recognized in anthropology: a survey to 175 evolutionary anthropologists (faculty members of graduate programs) finds that 78.7% of them regard cultural MLS as “important”, while 64.9% disagree with the statement “Group selection has no useful role to play in social science” (Yaworsky et al., 2015). Whether a similar acceptance rate of MLS by evolutionary biologists not working with humans is yet to be analyzed/surveyed. However, it is worth noticing that two recent analyses of biology, evolution, and behavior undergraduate textbooks show that MLS theory is generally dismissed as unimportant when compared to individual selection (Greene et al., 2025), with teaching on evolutionary transitions in individuality also being absent (La et al., 2025).

Our findings showing a marked increase in MLS research in the last 12 years (Figure 2), with 199 MLS studies for 2012–2024, indicates both that MLS is becoming more accepted as a conceptual framework and that many studies are using adequate sample sizes to ask questions across levels of biological organization. With the marked increase since 2012 and expanding acceptance of MLS as an conceptual evolutionary framework, many more ground-breaking studies are to come in the next few decades.

There are some caveats to our findings that the evidence for MLS is vast. First, we expect a publication bias towards studies finding positive outcomes, by which we mean that some studies where no selection at a higher level was found, were likely not captured. Despite this, our database does include studies in which higher-level selection or group properties were not important in explaining focal individual trait variation and fitness (Philson and Blumstein, 2023a; Philson and Blumstein, 2023b). Further, in several of the contextual analysis studies (Tsuji, 1995; Donohue, 2003; Weinig et al., 2007; Boege, 2010; Eldakar et al., 2010; Formica et al., 2011; Bolstad et al., 2012; Laiolo and Obeso, 2012; Fisher et al., 2017; Fisher et al., 2021) (Figure 4), the magnitude of selection was stronger at the individual than at the group level. Similarly, direct genetic effects are also usually stronger than indirect genetic effects, as shown by the meta-analysis of Santostefano et al. (2025) and through our database (but see Santostefano et al., 2021). However, because MLS should be evaluated in a case-by-case basis (Wilson and Wilson, 2007), this is not problematic for our framework: depending on the environmental context, case, and traits, it is expected that there will be cases in which there are no group effects or they are not as important as individual-level effects. It is, in fact, a central point that, unless a MLS perspective is applied, one would not know if or how strongly individual selection was causally related to trait evolution (Wade, 2026). Secondly, in order to have a distinct cut-off, we excluded MLS empirical evidence produced after 2024, thus missing new studies such as Lipowska et al. (2025), showing how bank vole holobionts selected for herbivorous capability evolved distinct and robust gut bacterial communities.

In general, we were quite strict in our search. For example, a study classically cited by some as the first MLS empirical study (Lewontin, 1962) was excluded, because, although it is based on real lab mice population data, the conclusions (about interdemic selection) are based on Monte Carlo simulations. Similarly, studies arguing that MLS is a "likely" (Dyer et al., 2005) or "possible" (Aviles, 1986) explanation were also excluded. Thus our total of 280 articles obtained is an underestimate of the evidence and conceptual use, because many more studies that clearly show results consistent with the MLS framework (Pope, 1992; Heinsohn and Packer, 1995; Ingvarsson, 2000; Papkou et al., 2023; Barnett et al., 2025), have historically avoiding using the term (Eldakar and Wilson, 2011; Greene et al., 2025). For example results based on Wright's fitness landscapes (Papkou et al., 2023) or on evolvability (Barnett et al., 2025), explicitly require a MLS perspective to understand them, but avoid the terminology. Although a MLS framework may not be explicitly mentioned by name, and in some cases may be avoided due to historical misconceptions (Eldakar and Wilson, 2011), it is implicit in experimental design and rationale.

5 Conclusions

In summary, a thorough search of the literature shows that contrary to common misconceptions which plagued the field since the 1960s, there is vast empirical evidence of selection acting at multiple levels and of the utility of assessing multilevel selection (MLS) both *in situ* and via experimental studies. We found 280 papers providing empirical support for MLS: 100 *in situ* and 180 laboratory experiments. The studies span many taxa and research methodologies, meaning MLS is not situational or an exception: MLS is a powerful evolutionary force in nature. Disregarding MLS will continue to hold the field of evolutionary biology back and prevent us from more fully understanding life on earth.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

CM: Writing – original draft, Writing – review & editing, Investigation, Data curation, Methodology, Visualization, Formal analysis, Conceptualization. AC: Formal analysis, Data curation, Writing – review & editing, Investigation. CP: Data curation, Investigation, Formal analysis, Writing – review & editing. OE: Writing – review & editing, Formal analysis, Data curation, Investigation. MW: Investigation, Writing – review & editing, Conceptualization, Formal analysis, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Funded by Fondecyt Regular Project No. 1240186 (ANID, Chile).

Acknowledgments

Many thanks to David Sloan Wilson for helpful comments and suggestions and to Mitchel Distin and the Multilevel Selection Initiative (https://www.prosocial.world/prosocial-initiatives/the-multilevel-selection-initiative), for support, helpful academic discussion, and prosociality. CM thanks ANID + Convocatoria Nacional Subvención a Instalación en la Academia Convocatoria Año 2021 + Folio no. SA77210019 and the Fondecyt Regular Project no. 1240186 (ANID, Convocatoria 2024). Many thanks to Felipe G. Serrano (https://illustrative-science.com/) for Figure 3 artwork.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author CM declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2026.1752597/full#supplementary-material

References

Aktipis, C. A., Boddy, A. M., Jansen, G., Hibner, U., Hochberg, M. E., Maley, C. C., et al. (2015). Cancer across the tree of life: cooperation and cheating in multicellularity. *Philos. Trans. R. Soc B* 370, 20140219. doi: 10.1098/rstb.2014.0219

Arnold, S. J., and Wade, M. J. (1984). On the measurement of natural and sexual selection: theory. *Evolution* 38, 709–719. doi: 10.2307/2408383

Aspi, J., Jäkäläniemi, A., Tuomi, J., and Siikamäki, P. (2003). Multilevel phenotypic selection on morphological characters in a metapopulation of Silene tatarica. *Evolution* 57 (3), 509–517.

Aviles, L. (1986). Sex-ratio bias and possible group selection in the social spider *Anelosimus eximius*. *Am. Nat.* 128, 1–12. doi: 10.1086/284535

Barnett, M., Meister, L., and Rainey, P. B. (2025). Experimental evolution of evolvability. *Science* 387, eadr2756. doi: 10.1126/science.adr2756

Baud, A., Casale, F. P., Barkley-Levenson, A. M., Farhadi, N., Montillot, C., Yalcin, B., et al. (2021). Dissecting indirect genetic effects from peers in laboratory mice. *Genome Biol.* 22, 216. doi: 10.1186/s13059-021-02415-x

Bijma, P. (2014). The quantitative genetics of indirect genetic effects: a selective review of modelling issues. *Heredity* 112, 61–69. doi: 10.1038/hdy.2013.15

Bijma, P., and Wade, M. J. (2008). The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* 21, 1175–1188. doi: 10.1111/j.1420-9101.2008.01550.x

Blackstone, N. W., Blackstone, S. R., and Berg, A. T. (2020). Variation and multilevel selection of SARS-CoV-2. *Evolution* 74, 2429–2434. doi: 10.1111/evo.14080

Boege, K. (2010). Induced responses to competition and herbivory: natural selection on multi-trait phenotypic plasticity. *Ecology* 91, 2628–2637. doi: 10.1890/09-0543.1

Bolstad, G. H., Pelabon, C., Larsen, L. K., Fleming, I. A., Viken, Å., and Rosenqvist, G. (2012). The effect of purging on sexually selected traits through antagonistic pleiotropy with survival. *Ecol. Evol.* 2, 1181–1194. doi: 10.1002/ece3.246

Bouwhuis, S., Vedder, O., Garroway, C. J., and Sheldon, B. C. (2015). Ecological causes of multilevel covariance between size and first-year survival in a wild bird population. *J. Anim. Ecol.* 84 (1), 208–218.

Bozdag, G. O., Zamani-Dahaj, S. A., Day, T. C., Kahn, P. C., Burnetti, A. J., Lac, D. T., et al. (2023). *De novo* evolution of macroscopic multicellularity. *Nature* 617, 747–754. doi: 10.1038/s41586-023-06052-1

Burch, J., Chin, M., Fontenot, B. E., Mandal, S., McKnight, T. D., Demuth, J. P., et al. (2024). Wright was right: leveraging old data and new methods to illustrate the critical role of epistasis in genetics and evolution. *Evolution* 78, 624–634. doi: 10.1093/evolut/qpae003

Buttery, N. J., Thompson, C. R. L., and Wolf, J. B. (2010). Complex genotype interactions influence social fitness during the developmental phase of the social amoeba *Dictyostelium discoideum*. *J. Evol. Biol.* 23, 1664–1671. doi: 10.1111/j.1420-9101.2010.02032.x

Costello, R. A., Cook, P. A., Brodie, E. D. III, and Formica, V. A. (2023). Multilevel selection on social network traits differs between sexes in experimental populations of forked fungus beetles. *Evolution* 77 (1), 289–303.

Craig, J. V., and Muir, W. M. (1996). Group selection for adaptation to multiple-hen cages: beak-related mortality, feathering, and body weight responses. *Poul. Sci.* 75, 294–302. doi: 10.3382/ps.0750294

Damuth, J., and Heisler, I. L. (1988). Alternative formulations of multilevel selection. *Biol. Philos.* 3, 407–430. doi: 10.1007/BF00647962

Darwin, C. (1871). *The Descent of Man* (London, United Kingdom: John Murray).

Donohue, K. (2003). The influence of neighbor relatedness on multilevel selection in the Great Lakes sea rocket. *Am. Nat.* 162, 77–92. doi: 10.1086/375299

Donohue, K. (2004). Density-dependent multilevel selection in the great lakes sea rocket. *Ecology* 85, 180–191. doi: 10.1890/02-0767

Dyer, K. A., Minhas, M. D., and Jaenike, J. (2005). Expression and modulation of embryonic male-killing in *Drosophila innubila*: opportunities for multilevel selection. *Evolution* 59, 838–848. doi: 10.1111/j.0014-3820.2005.tb01757.x

Eldakar, O. T., and Wilson, D. S. (2011). Eight criticisms not to make about group selection. *Evolution* 65, 1523–1526. doi: 10.1111/j.1558-5646.2011.01290.x

Eldakar, O. T., and Gallup, A. C. (2011). The group-level consequences of sexual conflict in multigroup populations. *PloS One* 6 (10), e26451.

Eldakar, O. T., Wilson, D. S., Dlugos, M. J., and Pepper, J. W. (2010). The role of multilevel selection in the evolution of sexual conflict in the water strider *Aquarius remigis*. *Evolution* 64, 3183–3189. doi: 10.1111/j.1558-5646.2010.01087.x

Fisher, D. N., Boutin, S., Dantzer, B., Humphries, M. M., Lane, J. E., and McAdam, A. G. (2017). Multilevel and sex-specific selection on competitive traits in North American red squirrels. *Evolution* 71, 1841–1854. doi: 10.1111/evo.13270

Fisher, D. N., LeGrice, R. J., and Painting, C. J. (2021). Social selection is density dependent but makes little contribution to total selection in New Zealand giraffe weevils. *Proc. R. Soc B: Biol. Sci.* 288, 20210696. doi: 10.1098/rspb.2021.0696

Formica, V. A., McGlothlin, J. W., Wood, C. W., Augat, M. E., Butterfield, R. E., Barnard, M. E., et al. (2011). Phenotypic assortment mediates the effect of social selection in a wild beetle population. *Evolution* 65, 2771–2781. doi: 10.1111/j.1558-5646.2011.01340.x

Francois, P., Fujiwara, T., and van Ypersele, T. (2018). The origins of human prosociality: Cultural group selection in the workplace and the laboratory. *Sci. Adv.* 4, eaat2201. doi: 10.1126/sciadv.aat2201

Frank, S. A. (2025). Natural selection at multiple scales. *Evolution* 79, 1166–1184. doi: 10.1093/evolut/qpaf037

Gitschlag, B. L., Tate, A. T., and Patel, M. R. (2020). Nutrient status shapes selfish mitochondrial genome dynamics across different levels of selection. *eLife* 9, e56686. doi: 10.7554/eLife.56686

Goodnight, C. (2013). On multilevel selection and kin selection: contextual analysis meets direct fitness. *Evolution* 67, 1539–1548. doi: 10.1111/j.1558-5646.2012.01821.x

Goodnight, C. J. (2015). Multilevel selection theory and evidence: a critique of Gardner, 2015. *J. Evol. Biol.* 28, 1734–1746. doi: 10.1111/jeb.12685

Goodnight, C. J., Schwartz, J. M., and Stevens, L. (1992). Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *Am. Nat.* 140, 743–761. doi: 10.1086/285438

Goodnight, C. J., and Stevens, L. (1997). Experimental studies of group selection: what do they tell us about group selection in nature? *Am. Nat.* 150, s59–s79. doi: 10.1086/286050

Gould, S. J., and Lloyd, E. A. (1999). Individuality and adaptation across levels of selection: how shall we name and generalize the unit of Darwinism? *Proc. Natl. Acad. Sci. U.S.A.* 96, 11904–11909. doi: 10.1073/pnas.96.21.11904

Greene, C. A., McPeek, S. J., Mitchem, L., Clark, A. B., Formica, V. A., Philson, C. S., et al. (2025). Multilevel selection in the margins: A review of its representation in undergraduate biology textbooks. *Ecol. Evol.* 15, e72493. doi: 10.1002/ece3.72493

Harms, K. E., Watson, D. M., Santiago-Rosario, L. Y. , and Mathews, S. (2023). Exposing the error hidden in plain sight: A critique of Calder's (1983) group selectionist seed-dispersal hypothesis for mistletoe "mimicry" of host plants. *Ecol. Evol.* 13, e10760. doi: 10.1002/ece3.10760

Heinsohn, R., and Packer, C. (1995). Complex cooperative strategies in group-territorial African lions. *Science* 269, 1260–1262. doi: 10.1126/science.7652573

Heisler, I. L., and Damuth, J. (1987). A method for analyzing selection in hierarchically structured populations. *Am. Nat.* 130, 582–602. doi: 10.1086/284732

Herron, M. D., and Michod, R. E. (2008). Evolution of complexity in the volvocine algae: transitions in individuality through Darwin's eye. *Evolution* 62, 436–451. doi: 10.1111/j.1558-5646.2007.00304.x

Hertler, S. C., Figueredo, A. J., and Peñaherrera-Aguirre, M. (2020). *Multilevel selection: Theoretical foundations, historical examples, and empirical evidence* (Berlin, Germany: Springer Nature).

Ingvarsson, P. K. (2000). Differential migration from high fitness demes in the shining fungus beetle, *Phalacrus substriatus*. *Evolution* 54, 297–301. doi: 10.1111/j.0014-3820.2000.tb00031.x

Jacquiod, S., Spor, A., Wei, S., Munkager, V., Bru, D., Sørensen, S. J., et al. (2022). Artificial selection of stable rhizosphere microbiota leads to heritable plant phenotype changes. *Ecol. Lett.* 25, 189–201. doi: 10.1111/ele.13916

Jany, J. L., and Pawlowska, T. E. (2010). Multinucleate spores contribute to evolutionary longevity of asexual Glomeromycota. *Am. Nat.* 175, 424–435. doi: 10.1086/650725

Jochum, M. D., McWilliams, K. L., Pierson, E. A., and Jo, Y. K. (2019). Host-mediated microbiome engineering (HMME) of drought tolerance in the wheat rhizosphere. *PloS One* 14, e0225933. doi: 10.1371/journal.pone.0225933

Johnson, N. C., Graham, J. H., and Smith, F. A. (1997). Functioning of mycorrhizal associations along the mutualism–parasitism continuum. *New Phytol.* 135, 575–585. doi: 10.1046/j.1469-8137.1997.00729.x

Johnson, N. C., and Marín, C. (2025). Functional team selection as a framework for local adaptation in plants and their belowground microbiomes. *ISME J.* 19, wraf137. doi: 10.1093/ismejo/wraf137

La, S., Grochau-Wright, Z. I., Hoskinson, J. S., Davison, D. R., and Michod, R. E. (2025). Translating research on evolutionary transitions into the teaching of hierarchical complexity in university biology courses. *Ecol. Evol.* 15, e72267. doi: 10.1002/ece3.72267

Laiolo, P., and Obeso, J. R. (2012). Multilevel selection and neighbourhood effects from individual to metapopulation in a wild passerine. *PloS One* 7, e38526. doi: 10.1371/journal.pone.0038526

Lande, R., and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* 37, 1210–1226. doi: 10.1111/j.1558-5646.1983.tb00236.x

Lau, J. A., and Lennon, J. T. (2012). Rapid responses of soil microorganisms improve plant fitness in novel environments. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14058–14062. doi: 10.1073/pnas.1202319109

Lee, I. P. A., Eldakar, O. T., Gogarten, J. P., and Andam, C. P. (2022). Bacterial cooperation through horizontal gene transfer. *Trends Ecol. Evol.* 37, 223–232. doi: 10.1016/j.tree.2021.11.006

Levins, R. (1970). "Extinction," in *Some Mathematical Questions in Biology*, vol. 2. Ed. M. Gerstenhaber (Providence, Rhode Island, US: American Mathematical Society), 77–107.

Lewontin, R. C. (1962). Interdeme selection controlling a polymorphism in the house mouse. *Am. Nat.* 96, 65–78. doi: 10.1086/282208

Linksvayer, T. A., Fondrk, M. K., and Page, R. E. Jr. (2009). Honeybee social regulatory networks are shaped by colony-level selection. *Am. Nat.* 173, E99–E107. doi: 10.1086/596527

Lipowska, M. M., Sadowska, E. T., Kohl, K. D., and Koteja, P. (2025). Experimental evolution of a mammalian holobiont: bank voles selected for herbivorous capability evolved distinct and robust gut bacterial communities. *ISME Commun.* 5, ycaf160. doi: 10.1093/ismeco/ycaf160

Lloyd, E. (2024). *Units and levels of selection, the Stanford Encyclopedia of Philosophy (Summer 2024 edition)*. Eds. E. N. Zalta and U. Nodelman Available online at: https://plato.stanford.edu/archives/sum2024/entries/selection-units/ (Accessed January 21, 2025).

Marín, C. (2015). Selección Multinivel: historia, modelos, debates, y principalmente, evidencias empíricas. *eVOLUCIÓN: Rev. la Sociedad Española Biología Evolutiva* 10, 51–70.

Marín, C. (2016). The levels of selection debate: taking into account existing empirical evidence. *Acta Biol. Colomb* 21, 467–472. doi: 10.15446/abc.v21n3.54596

Marín, C. (2021). Spatial and density-dependent multilevel selection on weed-infested maize. *Genet. Resour. Crop Evol.* 68 (3), 885–897.

Marín, C. (2024). Three types of units of selection. *Evolution* 78, 579–586. doi: 10.1093/evolut/qpad234

Marín, C., and Wade, M. J. (2025). Bring back the phenotype. *New Phytol.* 246, 2440–2445. doi: 10.1111/nph.70138

Maynard Smith, J. (1964). Group selection and kin selection. *Nature* 201, 1145–1146. doi: 10.1038/2011145a0

Montazeri, A., Mohammadi, S., M.Hesari, P., Ghaemi, M., Riazi, H., and Sheikhi-Mobarakeh, Z. (2023). Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. *Syst. Rev.* 12, 239. doi: 10.1186/s13643-023-02410-2

Moorad, J. A. (2013). Multi-level sexual selection: individual and family-level selection for mating success in a historical human population. *Evolution* 67, 1635–1648. doi: 10.1111/evo.12050

Mueller, U. G., Juenger, T. E., Kardish, M. R., Carlson, A. L., Burns, K. M., Edwards, J. A., et al. (2021). Artificial selection on microbiomes to breed microbiomes that confer salt tolerance to plants. *mSystems* 6, e01125–e01121. doi: 10.1128/mSystems.01125-21

Okasha, S. (2003). The concept of group heritability. *Biol. Philos.* 18, 445–461. doi: 10.1023/A:1024140123391

Okasha, S. (2006). *Evolution and the levels of selection* (New York, United States: Oxford University Press).

Papkou, A., Garcia-Pastor, L., Escudero, J. A., and Wagner, A. (2023). A rugged yet easily navigable fitness landscape. *Science* 382, eadh3860. doi: 10.1126/science.adh3860

Philson, C. S., and Blumstein, D. T. (2023a). Group social structure has limited impact on reproductive success in a wild mammal. *Behav. Ecol.* 34, 89–98. doi: 10.1093/beheco/arac102

Philson, C. S., and Blumstein, D. T. (2023b). Emergent social structure is typically not associated with survival in a facultatively social mammal. *Biol. Lett.* 19, 20220511. doi: 10.1098/rsbl.2022.0511

Philson, C. S., Martin, J. G., and Blumstein, D. T. (2025). Multilevel selection on individual and group social behaviour in the wild. *Proc. R. Soc B: Biol. Sci.* 292, 20243061. doi: 10.1098/rspb.2024.3061

Pope, T. R. (1992). The influence of dispersal patterns and mating system on genetic differentiation within and between populations of the red howler monkey (*Alouatta seniculus*). *Evolution* 46, 1112–1128. doi: 10.1111/j.1558-5646.1992.tb00623.x

Ratcliff, W. C., Denison, R. F., Borrello, M., and Travisano, M. (2012). Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1595–1600. doi: 10.1073/pnas.1115323109

Robinson, S. D., Schendel, V., Schroeder, C. I., Moen, S., Mueller, A., Walker, A. A., et al. (2023). Intra-colony venom diversity contributes to maintaining eusociality in a cooperatively breeding ant. *BMC Biol.* 21, 5. doi: 10.1186/s12915-022-01507-9

Sanchez, A., Bajic, D., Diaz-Colunga, J., Skwara, A., Vila, J. C., and Kuehn, S. (2023). The community-function landscape of microbial consortia. *Cell Syst.* 14, 122–134. doi: 10.1016/j.cels.2022.12.011

Sánchez, Á., Vila, J. C., Chang, C. Y., Diaz-Colunga, J., Estrela, S., and Rebolleda-Gomez, M. (2021). Directed evolution of microbial communities. *Annu. Rev. Biophys.* 50, 323–341. doi: 10.1146/annurev-biophys-101220-072829

Santostefano, F., Garant, D., Bergeron, P., Montiglio, P. O., and Reale, D. (2020). Social selection acts on behavior and body mass but does not contribute to the total selection differential in eastern chipmunks. *Evolution* 74 (1), 89–102.

Santostefano, F., Allegue, H., Garant, D., Bergeron, P., and Réale, D. (2021). Indirect genetic and environmental effects on behaviors, morphology, and life-history traits in a wild Eastern chipmunk population. *Evolution* 75, 1492–1512. doi: 10.1111/evo.14232

Santostefano, F., Moiron, M., Sánchez-Tójar, A., and Fisher, D. N. (2025). Indirect genetic effects increase the heritable variation available to selection and are largest for behaviors: a meta-analysis. *Evol. Lett.* 9, 89–104. doi: 10.1093/evlett/qrae051

Soltis, J., Boyd, R., and Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural group selection?: An empirical test. *Curr. Anthropol.* 36, 473–494. doi: 10.1086/204381

Stevens, L., Goodnight, C. J., and Kalisz, S. (1995). Multilevel selection in natural populations of *Impatiens capensis*. *Am. Nat.* 145, 513–526. doi: 10.1086/285753

Stinson, K. A., Brophy, C., and Connolly, J. (2011). Catching up on global change: new ragweed genotypes emerge in elevated CO2 conditions. *Ecosphere* 2 (4), 1–11.

Suárez, J., and Lloyd, E. A. (2023). *Units of selection* (Cambridge, United Kingdom: Cambridge University Press).

Swenson, W., Wilson, D. S., and Elias, R. (2000). Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9110–9114. doi: 10.1073/pnas.150237597

Takeuchi, N., and Kaneko, K. (2019). The origin of the central dogma through conflicting multilevel selection. *Proc. R. Soc B: Biol. Sci.* 286, 20191359. doi: 10.1098/rspb.2019.1359

Testard, C., Shergold, C., Acevedo-Ithier, A., Hart, J., Bernau, A., Negron-Del Valle, J. E., et al. (2024). Ecological disturbance alters the adaptive benefits of social ties. *Science* 384, 1330–1335. doi: 10.1126/science.adk0606

Tsuji, K. (1995). Reproductive conflicts and levels of selection in the ant *Pristomyrmex pungens*: contextual analysis and partitioning of covariance. *Am. Nat.* 146, 586–607. doi: 10.1086/285816

Turchin, P., and Gavrilets, S. (2021). Tempo and mode in cultural macroevolution. *Evol. Psychol.* 19, 4. doi: 10.1177/14747049211066600

Volis, S., Mendlinger, S., and Ward, D. (2002). Differentiation in populations of Hordeum spontaneum along a gradient of environmental productivity and predictability: life history and local adaptation. *Biol. J. Linn. Soc.* 77 (4), 479–490.

Wade, M. J. (1976). Group selections among laboratory populations of Tribolium. *Proc. Natl. Acad. Sci. U.S.A.* 73, 4604–4607. doi: 10.1073/pnas.73.12.4604

Wade, M. J. (2026). "Sewall wright and the shifting balance theory," in *Encyclopedia of Evolutionary Biology*, *2nd ed.* Eds. J. B. Wolf and C. A. De Moraes Russo (Academic Press, Amsterdam, the Netherlands), 387–395. doi: 10.1016/B978-0-443-15750-9.00111-7

Weinig, C., Johnston, J. A., Willis, C. G., and Maloof, J. N. (2007). Antagonistic multilevel selection on size and architecture in variable density settings. *Evolution* 61, 58–67. doi: 10.1111/j.1558-5646.2007.00005.x

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought* (Princeton, United States: Princeton University Press).

Wilson, D. S. (1975). A theory of group selection. *Proc. Natl. Acad. Sci. U.S.A.* 72, 143–146. doi: 10.1073/pnas.72.1.143

Wilson, D. S., Madhavan, G., Gelfand, M. J., Hayes, S. C., Atkins, P. W., and Colwell, R. R. (2023). Multilevel cultural evolution: From new theory to practical applications. *Proc. Natl. Acad. Sci. U.S.A.* 120, e2218222120. doi: 10.1073/pnas.2218222120

Wilson, D. S., Philip, M. M., MacDonald, I. F., Atkins, P. W., and Kniffin, K. M. (2020). Core design principles for nurturing organization-level selection. *Sci. Rep.* 10, 13989. doi: 10.1038/s41598-020-70632-8

Wilson, D. S., and Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behav. Brain Sci.* 17, 585–608. doi: 10.1017/S0140525X00036104

Wilson, D. S., and Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* 82, 327–348. doi: 10.1086/522809

Wright, S. (1945). Tempo and mode in evolution: a critical review. *Ecology* 26, 415–419. doi: 10.2307/1931666

Yaworsky, W., Horowitz, M., and Kickham, K. (2015). Gender and politics among anthropologists in the units of selection debate. *Biol. Theory* 10, 145–155. doi: 10.1007/s13752-014-0196-5

Yu, S. R., Zhang, Y. Y., and Zhang, Q. G. (2023). The effectiveness of artificial microbial community selection: a conceptual framework and a meta-analysis. *Front. Microbiol.* 14. doi: 10.3389/fmicb.2023.1257935

Zhu, Y. H., Weiner, J., Jin, Y., Yu, M. X., and Li, F. M. (2022). Biomass allocation responses to root interactions in wheat cultivars support predictions of crop evolutionary ecology theory. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.858636

Zhu, Y. H., Weiner, J., and Li, F. M. (2019a). Root proliferation in response to neighbouring roots in wheat (*Triticum aestivum*). *Basic Appl. Ecol.* 39, 10–14. doi: 10.1016/j.baae.2019.07.001

Zhu, Y. H., Weiner, J., Yu, M. X., and Li, F. M. (2019b). Evolutionary agroecology: Trends in root architecture during wheat breeding. *Evol. Appl.* 12, 733–743. doi: 10.1111/eva.12749

Appendix 1. Supplementary Discussion: Multilevel Selection is not equivalent to inclusive fitness.

Appendix from: Marín C, Clark AB, Philson CS, Eldakar OT, Wade MJ. 2025. Abundant empirical evidence of multilevel selection revealed by a bibliometric review.

> “*Inclusive fitness theory, summarised in Hamilton’s rule, is a dominant explanation for the evolution of social behaviour. A parallel thread of evolutionary theory holds that selection between groups is also a candidate explanation for social evolution. The mathematical equivalence of these two approaches has long been known*”. Marshall (2011; p. 326).

Many authors, including Okasha (2006, 2016), Marshall (2011), Ågren (2021), and others, have asserted that the mathematical theories of inclusive fitness (IF) and multilevel selection (MLS) are equivalent, a claim whose veracity depends upon one’s definition of equivalence. The accepted definition of ‘equivalence’ appears to be that both theories give rise to Hamilton’s Rule, $rb - c > 0$, the condition for the evolution of costly cooperation. This interpretation of ‘equivalence’ is supported by the modifications of Hamilton’s Rule that are added when selection is strong and/or fitnesses are non-additive to permit predicting the *direction* of allele frequency change. Here, we use the simplest haploid model to illustrate four ways in which the MLS and IF theories are *not* equivalent, even though both predict the direction of selection.

Hamilton (1963, 1964a, b) defined the inclusive fitness, $w_{IF,i}$, of individual i, to be the sum of two components: (1) the direct genetic effect of its own genes on its own fitness *plus* (2) the indirect effect of individual i on the fitness of others weighted by r, the genetic relatedness between individual i and those others. When applied to the evolution of altruism between interacting pairs, Hamilton’s direct effect, c, was defined as the increment of fitness cost to the performer of an altruistic act ($c < 0$), while b was defined as the increment of fitness benefit accrued by another individual, the recipient of the altruistic act. We will use the notation W for classical fitness used in MLS theory and W_{IF} for inclusive fitness used in IF theory.

Allow a single allele, A, in frequency, p, to determine the altruistic phenotype and its alternative allele, a, in frequency, q = (1 – p), be the non-altruist phenotype. There are three possible kinds of pairwise genetic interactions: (1) those between two altruists, which occur with frequency (p2 + rpq); (2) those between an altruist and a non-altruist, which occur with frequency 2(pq)(1 – r); and, (3) those between two non-altruists, which occur in frequency (q2 + pqr). We will first develop the classical fitness approach used in MLS theory and then the inclusive fitness approach of IF Theory. Lastly, we will contrast the properties of the two approaches to illustrate the dimensions in which they are or are not equivalent even in the simplest genetic model.

Classical Fitness Approach of MLS Theory.

In MLS theory, fitnesses of altruists and non-altruists depend upon group composition. Let’s imagine three groups (of two individuals each) with different compositions. In group 1, both altruists have fitness $W_A = 1 + b - c$, so this group’s mean fitness, $W_{AA} = 1 + b - c$. In group 2, the fitnesses of the altruist and non-altruist are $W_A = 1 – c$, $W_a = 1 + b$, for a mean fitness, $W_{Aa} = 1 + (b - c)/2$. In group 3, both are non-altruists, each with fitness $W_a = 1$, so W_{aa} is also 1. Mean classical fitness, W, averaged across the distribution of interacting pair frequencies, equals:

$$W = 1 + p(b – c). \tag{1}$$

From eq. (1), it is clear that genetic relatedness between interactants does not affect W, the total amount of incremental fitness costs and benefits. However, relatedness redistributes the benefits toward altruists (because A x A interactions are increased in frequency by relatedness, $r > 0$) and away from non-altruists (because A x a interactions are reduced in frequency).

If pairs are randomly assembled ($r = 0$), then, a fraction, p, of an altruist's interactions, are with another altruist and, in that fraction of cases, the focal altruist receives a beneficial indirect effect, while a fraction, $(1 – p)$, are with a non-altruist and the focal altruist receives nothing. Hence, the classical individual fitness of an altruist is: $w_A = 1 – c + pb$, where p is the frequency of other altruists encountered by the focal altruist as given by eq. (1). The same process can be applied to the non-altruist, so its fitness is: $w_a = 1 + pb$. Owing to the cost to the behavior, $w_A < w_a$ and for randomly pairing individuals, altruists are always selected against.

However, when pairs are non-randomly associated and *genetically related* to one another (i.e., $r > 0$), the fitness of altruists is increased, $w_A = 1 – c + (p + rq)b$, and the fitness of non-altruists, $w_a = 1 + p(1 – r)b$, is decreased. In this way, mean classical fitness, averaged across the two genotypes is: $W = pw_A + qw_a = 1 + p(b – c)$, i.e., the same quantity as mean fitness averaged across interacting pairs. Altruist fitness exceeds mean fitness, when $(w_A – W) = q(rb – c) > 0$. Thus, altruists have a fitness greater than the population mean (and so a *relative fitness* greater than 1) whenever Hamilton's Rule, $rb > c$, is true, and, conversely, when it is false. The fitness increment of the non-altruist is the difference, $(w_a – W) = -p(rb – c)$; this is *less than zero* when Hamilton's Rule is true.

Mean fitness changes when the genotype frequencies of altruists and non-altruists change. The change in the frequency of altruists, Δp, equals:

$$\Delta p = (p/W)(w_A – W) \tag{2a}$$
$$\Delta p = (pq/W)(rb – c) \tag{2b}$$

So far, this is standard population genetics for a haploid altruism model. One can also calculate the average gene frequency change by individual selection, $\Delta p_{Individual}$, which occurs *within interacting groups*, by treating each pairing as though it were a separate 'population' subject to the rules of gene frequency change just applied above. This calculation is simplified because there is no genetic variation within interaction group types 1 and 3, so there cannot be individual selection within either group. This means that $\Delta p_{Individual} = -\{(b + c)/2W\}pq(1 – r)$, the change in altruist frequency within group 2 interactions $(-[b + c]/4W_{Aa})$, multiplied by the frequency of those interactions after selection, $2pq(1 – r)(W_{Aa}/W)$, where (W_{Aa}/W) is the relative fitness of type 2 interaction groups. Note that $\Delta p_{Individual}$ is negative: on average, altruists have lower fitness within interacting pairs and their frequency within them declines.

There is also a component of gene frequency change owing to group selection, Δp_{Group}, because mean group fitness covaries with group altruist frequency. This component of selection equals: $\Delta p_{Group} = +\{(b - c)/W\}(pq[1 + r]/2)$. Their sum, $\Delta p_{Individual} + \Delta p_{Group}$, equals the total Δp given in eq. (2b) (population genetic models of altruism can always be partitioned into within and between group components) (Wade 1980; Wade 1982; Wade and Breden 1987).

The Inclusive Fitness Approach.

For this model, Hamilton defines the inclusive fitness of altruists, as: $w_{IF,A} = 1 - c + rb$, and that of non-altruists, as: $w_{IF,a} = 1$. Mean inclusive fitness, is defined as: $W_{IF} = p\, w_{IF,A} + q\, w_{IF,a} = 1 + p(rb - c)$. Note that $(w_{IF,A} - W_{IF}) = +q(rb - c)$ and that $(w_a - W) = (w_{IF,a} - W_{IF}) = -p(rb - c)$. Clearly, whenever $(w_{IF,A} - W_{IF}) > 0$, the altruist allele will increase in frequency and when $(w_{IF,A} - W_{IF}) < 0$, it will decrease in frequency. Hamilton's Rule, $(rb - c) > 0$, is the condition that permits the evolution of altruistic behavior. And, we arrived at this conclusion much more quickly and simply that we did in the classical case, especially when we broke the classical case down into opposing components, $\Delta p_{Individual}$ and Δp_{Group}. If getting to the condition for positive evolution of genetic altruism is what is meant by 'equivalence', MLS and IF theories are equivalent and IF theory appears to be much easier to apply.

The Ways in which MLS and IF Theories are NOT Equivalent.

First, note that IF theory does not predict gene frequency change because, although $(w_{IF,A} - W_{IF}) = (w_A - W)$, the change in allele frequency equals:

$$\Delta p \neq (pq/W_{IF})(w_{IF,A} - W_{IF}), \qquad (3)$$

This is owing to the fact that:

$$W_{IF} = 1 + p(rb - c) \neq W = 1 + p(b - c). \qquad (4)$$

Whenever Hamilton's Rule is true, then necessarily, $W > W_{IF}$. That is, when $(rb - c) > 0$, $(b - c) > (rb - c)$, because $rb < b$. If one were to use W_{IF} in the denominator of Δp, the rate of evolutionary change would be over-estimated when $(rb - c) > 0$, and underestimated when $(rb - c) < 0$. **So, MLS and IF theories both predict the correct *direction* of evolutionary change, but only MLS predicts the correct *magnitude* of that change.**

Classical mean absolute fitness, W, predicts change in population size, because $N_{t+1} = W_t N_t$. And, W is necessary for calculating *relative fitness*, which keeps allele frequencies summing to 1 *after* they are changed by selection. That is, the sum of the allele frequency distribution before selection, $p + q$, and after selection, $p' + q' = p(w_A/W) + q(w_a/W)$, sum to 1.

The change in classical mean fitness, ΔW, and the change in inclusive fitness by selection, ΔW_{IF}, are obtained by taking the difference between the respective values after and before selection.

$$\Delta W = (pq/W)(rb - c)(b - c) \qquad (5a)$$
$$\Delta W_{IF} = (pq/W)(rb - c)^2 \qquad (5b)$$

Note first, that, because W is necessary for calculating Δp, it is also necessary here for calculating ΔW_{IF}. In this way, IF theory is *dependent upon* quantities calculated from the MLS theory fitness definition. Secondly, note that: $\Delta W \neq \Delta W_{IF}$. Thirdly, note that ΔW_{IF} is always positive while ΔW can be negative because $(rb - c)$ can be negative. Positive ΔW_{IF} does not mean that 'altruists are always favored'. Differently put, maximizing W_{IF} when Hamilton's Rule is not met, means eliminating altruists. However, whenever Hamilton's Rule is true, $\Delta W > \Delta W_{IF}$, because whenever $(rb - c)$ is positive, so is $(b - c)$, because $rb < b$. ***Mean classical fitness is always larger than and is always changing faster than mean inclusive fitness whenever Hamilton's Rule holds.***

Bjima (2010) showed how these results can be connected to Fisher's Fundamental Theorem. Remembering that $(w_A - W) = (w_{IF,A} - W_{IF}) = q(rb - c)$ and that $(w_a - W) = (w_{IF,a} - W_{IF}) = -p(rb - c)$, the *additive genetic variance* in individual classical fitness, $\sigma^2(w)$, equals the additive variance in inclusive fitness, $\sigma^2(w_{IF})$:

$$\sigma^2(w) = \sigma^2(w_{IF}) \tag{6a}$$
$$\sigma^2(w) = pq(rb - c)^2 = W\Delta W_{IF} \tag{6b}$$

Fisher's Fundamental Theorem*, which states that the increase in mean fitness is proportional to the additive genetic variance in fitness****, describes the change in mean inclusive fitness,*** ΔW_{IF}***, but NOT the change in mean fitness,*** ΔW***, when there are indirect genetic effects***. However, as first pointed out by Bijma (2010), it appears that **inclusive fitness extends Fisher's definition of individual breeding value, *to include indirect genetic effects that a selected individual perpetrates upon its relatives***. Because the change in inclusive fitness is always positive, individuals with adaptations for increasing mean inclusive fitness are favored in this context by individual selection, even though the full effects on the second term of the Price Equation are ***not*** captured by eqs. (6a) and (6b); i.e., the extended definition of individual breeding value to include indirect genetic effects perpetrated upon relatives does not describe ΔW.

It is unfortunate that, even in the simplest haploid model, IF theory lacks utility in predicting change in population size, change in total mean fitness, change in gene frequency, or change in linkage disequilibrium by selection in more complicated, multilocus models (algebra not shown). These differences in the properties and utility of these two different definitions of fitness are evidence that the two approaches ***are not equivalent*** in many of the ways that one might desire in a general genetic theory of evolution. Getting the direction of allele frequency change correctly is important, but ultimately it is only one dimension of equivalence.

Moreover, group selection can add a component of group fitness *beyond that embodied by the indirect fitness effects perpetrated by individuals upon their conspecific relatives*. That is essentially what animal and plant breeders do to mitigate the deleterious competitive fitness effects that high-yielding individuals have on the yield of their unfortunate neighbors. Differently put, competitive indirect fitness effects on neighboring relatives represent an impediment to maximizing mean fitness in natural populations and yield in domesticated animals and plants. Because of the group selection imposed by the breeder, these deleterious competitive effects on fitness are not an impediment to artificial selection and the genetic response to that selection. In natural populations, there are also agents of group selection that, like breeders, affect individual fitnesses differently from and in addition to the indirect fitness effects perpetrated by individuals on their conspecific relatives, which are captured by the concept of inclusive fitness (see Bijma and Wade 2008, for a more detailed exposition).

References

- Ågren JA. 2021. *The gene's-eye view of evolution*. New York, United States: Oxford University Press.
- Bijma P. 2010. Fisher’s fundamental theorem of inclusive fitness and the change in fitness due to natural selection when conspecifics interact. *J. Theor. Biol.* **23**, 194-206. doi: 10.1111/j.1420-9101.2009.01895.x
- Bijma P, Wade MJ. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* **21**, 1175-1188. doi: 10.1111/j.1420-9101.2008.01550.x
- Hamilton WD. 1963. The evolution of altruistic behavior. *Am. Nat.* **97**, 354-356. doi: 10.1086/497114
- Hamilton WD. 1964b. The genetical evolution of social behaviour. II. *J. Theor. Biol.* **7**, 17-52. doi: 10.1016/0022-5193(64)90039-6
- Hamilton WD. 1964a. The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1-16. doi: 10.1016/0022-5193(64)90038-4
- Marshall JA. 2011. Group selection and kin selection: formally equivalent approaches. *Trends. Ecol. Evol.* **26**, 325-332. doi: 10.1016/j.tree.2011.04.008
- Okasha S. 2016. The relation between kin and multilevel selection: an approach using causal graphs. *Br. J. Philos. Sci.* **67**, 435-470. doi: 10.1093/bjps/axu047
- Okasha S. 2006. *Evolution and the levels of selection*. New York, United States: Oxford University Press.
- Wade MJ. 1980. Kin selection: its components. *Science* **210**, 665-667. doi: 10.1126/science.210.4470.665
- Wade MJ. 1982. The effect of multiple inseminations on the evolution of social behaviors in diploid and haplo-diploid organisms. *J. Theor. Biol.* **95**, 351-368. doi: 10.1016/0022-5193(82)90250-8
- Wade MJ, Breden FJ. 1987. Kin Selection in complex groups: Mating structure, migration structure, and the evolution of social behaviors. In: Chepko-Sade BD, Halpin ZJ (Eds). M*ammalian Dispersal Patterns: the Effects of Social Structure on Population Genetics*. Chicago, University of Chicago Press. Pp. 273-283.

Appendix 2. Supplementary Methods.

Appendix from: Marín C, Clark AB, Philson CS, Eldakar OT, Wade MJ. 2025. Abundant empirical evidence of multilevel selection revealed by a bibliometric review.

The bibliometric database generated in this article is available as: **Appendix 3.**

Items on this Supplementary Methods file:
- Supplementary Table S1 (p. 2).
- Supplementary Methods (p. 4-15).
- Supplementary Table S2 (p. 4).
- Supplementary Table S3 (p. 16).
- References (p. 18).

Supplementary Table S1. BIBLIO check-list (Montazeri et al. 2023) for the bibliometric review of the article “Abundant empirical evidence of multilevel selection revealed by a bibliometric review”.

Section/topic	**Item no**	**Check-list item**	**Reported on page no**
Title			
Identification	1	Identify the report as a bibliometric review in the title	1
Issues/topics	2	Indicate the key issues/topics under investigation and coverage of time period	1
Abstract			
Structured summary	3	Structured summary including (as applicable): background, methods, results (key findings), and conclusions	1
Introduction/background			
Justification/rationale/explanation	4	Present review of existing knowledge and epidemiological information	2, 3
Objectives	5	Statement of the objective(s) or question(s)	2, 3
Methods			
Search engines (data sources)	6	Describe all information sources (such as electronic databases, contact with study authors, trial registers, or other grey literature sources)	3
Search strategy	7	Keywords and systematization criteria (date of search, language, type of document) for the search	3
Time period	8	The period that the review covers and the justification	3
Eligibility criteria	9	Describe all inclusion and exclusion criteria, languages, study design, type of publication, and time period	3-5 & this file
Data refinement (data selection procedure)	10	Remove the irrelevant articles; inspection to eliminate duplicate and unrelated articles (after evaluation of the title, abstract, and content)	3-4, & this file
Quality assessment (optional)	11	Assessment of papers by three authors and the use of assessing check-lists	
Data synthesis	12	Describe the methods used for summarizing, handling, synthesis, tabulations, or schematic displays. Describe how the data were analysed	5, & this file
Results			
Descriptive findings (statistics)	13	- Provide details of the search and selection process in a flow diagram - Number of citations retrieved (number of publication, year of publication, type of documents, country of publication, articles with the highest impact, most impactful authors, most impactful articles, authors with the highest production, top journals, top institutions, …)	5, 6 & Figure 1
Schematic map and trend	14	Summarize and/or present the schematic maps and trends using an appropriate software to present citations, journals, authors, top journals, time trends, emerging literature, and any relevant indicators (as applicable)	Not applicable
Tabulation and summarizing the findings	15	General recommendation: Studies under consideration could be summarized and organized by different subtitles and different scenarios. Regardless, results need to be presented in separate tables covering each subtitle. The followings are some options that could help to summarize the findings Option 4: - Start the presentation with a historical view [when and who first published on the topic] - Report on review papers. The result should be listed in a separate table,	5, 6

Section/topic	**Item no**	**Check-list item**	**Reported on page no**
		and also specify the review type (scoping review, narrative review, systematic review, and meta-analysis) - Summarize the findings according to different subtitles relevant to the main topic	
Synthesis of findings	16	Synthesize the findings as much as possible, find the gap, and propose a model, hypothesis, etc. (if applicable)	Not applicable
Discussion			
Summary of evidence	17	Summarize the main findings. The findings should be presented in more “general” or “accessible” terms	7
Interpretation	18	Include interpretation consistent with results. Explanations for observed outcomes, similarities, and differences reported would be essential	7, 9, 10
Strengths and limitations	19	Discuss the strengths and limitations	9, 10
Conclusion(s)	20	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications	10

Supplementary Methods

General exclusion/inclusion criteria and details for each category and sub-category of multilevel selection empirical studies.

This bibliometric review considers only empirical studies of multilevel selection (MLS), by which we mean studies that used either *in situ* observations/measurements (‘MLS *in situ*’) and/or lab experiments under controlled conditions (‘Experimental’). All studies employing mathematical modelling were excluded. Specifically, **MLS *in situ*** studies included any study done in the field, in natural conditions, without controlled conditions. They included measurements of unaltered natural biological populations or communities and *in situ* manipulations, alterations, or experiments done to natural biological populations or communities. This also included crops grown outside the lab/greenhouse, and farm animals when analysing population datasets of unaltered or non-experimentally-manipulated populations. MLS **Experimental** studies included any study done under laboratory-controlled conditions and/or subjected to experimental manipulations, including plants growing in greenhouse conditions and farm animals when experimental manipulations were done.

Based in our general definition of MLS (Box 1): “MLS occurs when natural selection simultaneously acts at two or more different levels in a nested biological hierarchy (Damuth and Heisler 1988; Okasha 2006; Wilson and Wilson 2007; Marín 2024). Specifically, MLS occurs when there is differential reproduction of groups in addition to reproduction of individual entities within them, or when the differential reproduction of individuals is based on their group composition or characteristics (like the social environment)”, we only included studies that either show differential reproduction of entire groups of individuals or that show effects of group composition on individual fitness – in other words effects of the social environment on individual reproduction.

In order to systematically analyse all studies, we divided the **MLS *in situ*** studies into five categories, and 31 sub-categories. MLS **Experimental** studies were divided into four categories, and 39 sub-categories. Thus, the 280 papers analysed in this study were divided in a total of 70 sub-categories. Below, we explain each category (**in bold**) and sub-category in detail (underlined), providing the number of papers and ID of the papers per each sub-category (using []).

Supplementary Table S2. Categories (9) and sub-categories (70) of topics/methods/assessments of multilevel selection (MLS) *in situ* and in Experimental studies. The number of papers per category/sub-category are presented in parenthesis. Total number of papers: 280.

MLS *in situ*
1. Cultural multilevel selection (No. of papers = 10; 1995 – 2023).
2. Dataset analyses (No. of papers = 18; 1987 – 2024): Molecular sequencing - Microsatellites (1); Molecular sequencing - Gene Transfer Agents (2); Molecular sequencing - Fingerprinting (1); Molecular sequencing - Colonies’ microsatellites (1); Molecular sequencing - Colonies’ genome-wide association studies (1); Molecular sequencing - Clones’ single nucleotide polymorphisms (1); Molecular sequencing - Alleles for cooperation (1); Populations (2); Microbiomes (1); Hybrid zone (2); Communities (1); Phylogenetics - Price equation (1); Phylogenetics - Heritability (1); Phylogenetics (2).
3. Indirect Genetic Effects (IGE) (No. papers = 29; 2008 – 2023): Wild animals (7); Farm animals (22).
4. Group effects (No. papers = 18; 1994 – 2024): Group heritability - Wild animals (1); Group heritability - Farm animals (1); Social networks – Wild animals (4); Selection – Wild animals (4); Group effects – Wild animals (1); Colony selection - Trait variation (1); Colony selection (1); Colony selection - Personality (1); Field experiment - Wild plants (1); Field experiment – Wild animals (2); Field experiment – Crops (1).

5. Contextual analysis (No. papers = 25; 1995 – 2023): Unaltered – Humans (1); Unaltered (11); Manipulated (13).

Experimental MLS

1. Lab experiment (No. papers = 107; 1976 – 2024): Population heritability – Wild animals (1); Molecular sequencing – Virus' RNA (1); Molecular sequencing – Selfish genetic elements (1); Molecular sequencing – Colonies' microsatellites (1); Molecular sequencing – Alleles (1); Microbiome assessment – Wild animals (1); Group selection treatments – Wild plants (2); Group selection treatments – Wild animals (18); Group selection treatments – Virus (1); Group personality – Wild animals (1); Group fitness – Wild animals (1); Group effects (Social networks) – Wild animals (6); Group effects – Wild animals (14); Group effects – Farm animals (3); Group effects – Crops (2); Culture experiment – Fungi (1); Culture experiment – Bacteria (5); Culture experiment – Algae (1); Community selection treatments – Wild animals (2); Community heritability – Wild animals (1); Colony trait variation – Wild animals (8); Colony selection – Wild animals (8); Colony fitness – Wild animals (2); Clonal lineages – Fungi (1); Clonal lineages – Algae (1); IGE Experiment – Wild animals (14); IGE Experiment – Plants (1); IGE Experiment – Microorganisms (1); IGE Experiment - Lab mice (1); IGE Experiment – Farm animals (6).
2. Breeding through group selection (No. papers = 32; 1996 – 2023): Breeding through group selection (21); IGE Breeding – Wild animals (1); IGE Breeding - Farm animals (9); IGE Breeding - Crops (1).
3. Psychology experiment (No. papers = 19; 2011 – 2024).
4. Artificial selection (No. papers = 22; 2000 – 2023): Single species (4); Multicellularity (12); Microbiome (5); Consortia (1).

MLS *in situ* studies

1. Cultural multilevel selection (No. of papers = 10; 1995 – 2023)

- Description: Studies that implement the MLS definition (see above) but for cultural traits; thus, traits that confer a group-level advantage can spread via cultural group selection [174]. Several of these studies use analyse small-scale societies datasets through the criteria of different group/multilevel selection models regarding group formation/extinction [23], between-group competition [183], cultural adaptation [147], and kinship/clan formation – the latter analysed through the Standard Cross-Cultural Sample database [241]. One study sketches the evidence on how cultural multilevel selection explaining human cooperation [148], while other study [237] analyses a database of past societies history (Seshat: Global History Databank), reaching similar conclusions regarding the importance of cultural multilevel selection. An analysis of a third database (Equinox 2020) was analysed through a MLS framework [264] to dissect the interactions of between-group competition, socio-political complexity, and differentiation-integration effects. The remaining studies [174, 256, 259] assessed cultural multilevel selection in modern-day human populations or organizations.
- Papers ID's: 23, 147, 148, 174, 183, 237, 241, 256, 259, 264.

2. Dataset analyses (No. of papers = 18; 1987 – 2024)

- Description: These studies analysed historical, published, or experimental data (molecular and phenotypic) to infer MLS processes occurring in natural populations or communities. Three main major sub-categories involve this type of studies: *Molecular sequencing*, *Dataset analyses*, and *Phylogenetics* analyses. Each is described below.

Molecular sequencing (No. of papers = 8; 2009 – 2024): Studies that implemented any sort of molecular sequencing to natural populations, using different tools, from single-nucleotide polymorphism analysis to a genome-wide association study. Each sub-category is described below.

→ <u>Molecular sequencing – Microsatellites (1 paper)</u>: The authors investigated gene flow between wild boar and pig demes, under the framework of Wright's Shifting balance theory (Wade 2026). They analysed the *MC1R* gene (common in pigs) and microsatellite neutral loci in 62 pregnant wild boars as markers of hybridization. The authors found evidence of gene flow: regardless of body size and phyletic effects, wild boar sows bearing nonsynonymous MC1R mutations produced larger litters
- Paper ID: 142.
→ <u>Molecular sequencing – Gene Transfer Agents (2 papers)</u>: Both of these studies show that bacterial gene transfer agents (GTA's), which are enigmatic virus-like elements integrated into bacterial genomes, evolve through multilevel selection. Such GTA's can deliver DNA to neighbouring, closely related bacterial cells. GTA's are particularly functionally important during stressful conditions, on which they support stress response, DNA repair, and biofilm formation.
- Papers ID's: 209, 274.
→ <u>Molecular sequencing – Fingerprinting (1 paper)</u>: The authors used DNA fingerprinting to show that gene flow among nests, as well as spatial structuring at the level of regions, local populations, and nests, were compatible with interdemic selection in the social spider *Stegodyphus dumicola*.
- Paper ID: 67.
→ <u>Molecular sequencing – Colonies' microsatellites (1 paper)</u>: This study genotyped the queens of 23 field colonies of *Reticulitermes speratus*. The authors found that clone variation in the queen population reduces as colonies develop. The results show that while the production of earlier and more parthenogenetic eggs is advantageous for winning the competition for personal reproduction, it comes at a great cost to the colony. This study shows the interactions between individual level and colony-level selection on parthenogenesis by termite queens.
- Paper ID: 280.
→ <u>Molecular sequencing – Colonies' GWAS (1 paper)</u>: This study used genome-wide association studies (GWAS) to determine the influence of individual vs. group genotypes on aggression in honey bees. The authors found that individual behaviour is influenced by colony environment, which in turn, is shaped by allele frequency within colonies. The authors conclude that they "found links between colony genetics and group behaviour and also, molecular evidence for group-level selection, acting at the colony level".
- Paper ID: 203.
→ <u>Molecular sequencing – Clones' SNP's (1 paper)</u>: This study implemented single nucleotide polymorphisms (SNP's) analysis to show that branching events in the sea grass *Zostera marina* clones or genets, lead to population bottlenecks of tissue that result in the evolution of genetically differentiated ramets in a process of somatic genetic drift. The authors found that thousands of SNP's segregated among ramets. This study provides "evidence for multiple levels of selection during the evolution of sea grass genets".
- Paper ID: 217.
→ <u>Molecular sequencing – Alleles for cooperation (1 paper)</u>: The authors implemented a game-theoretical approach, empirically tested in *Arabidopsis thaliana*, where they found a major effect locus where the rarer allele was associated with increased cooperation and productivity in high-density stands.
- Paper ID: 248.

Dataset analyses (No. papers = 6; 1996 – 2022): These studies analysed historical or published data to infer MLS processes occurring in natural populations or communities.

→ Dataset analyses - Populations (2 papers): Both of these studies were conducted in social spiders, finding that the colonies of such spiders reproduce by fission [39] and significant gene flow among established colonies [34].
- Papers ID's: 34, 39.

→ Dataset analyses - Microbiomes (1 paper): This study analysed a total of 39 experiments published in the last 20 years that measured the asexual survival and reproduction of strains as a function of their initial frequency, holding constant the total number of individuals. The authors found that "strain and multilevel fitness outcomes were both effective for quantitatively comparing social selection in different datasets".
- Paper ID: 235.

→ Dataset analyses - Hybrid zone (2 papers): Both of these studies were conducted with butterflies form the genus Heliconius. The hybrid zone of Heliconius has rapidly moved in northern South America over the last few decades, which, according to these studies, represents strong evidence of phase III (interdemic selection) of Wright's Shifting Balance Theory (Wade 2026).
- Papers IDs: 47, 72.

→ Dataset analyses - Communities (1 paper): This study found evidence of MLS at the community level by analysing three years of common garden data for 102 species comprising arthropod communities, associated with nine clonally replicated *Populus angustifolia* genotypes. The authors estimated broad-sense community heritability.
- Paper ID: 246.

Phylogenetics (No. papers = 4; 1987 – 2015): These studies implemented phylogenetic analyses either to assess selection at the species level or to explain the evolution of complexity/multicellularity across an algae phylogenetic tree.

→ Phylogenetics - Price equation (1 paper): This study used an extension of the Price equation (Price 1970) to quantify the mechanisms of body size macroevolution in mammals from the latest Palaeocene and earliest Eocene of the Bighorn and Clarks Fork Basins of Wyoming, United States. The Price equation (Price 1970) was used to used to separate directional macroevolutionary change into components attributable to species selection and anagenetic change. Through this method, the authors find strong evidence of species selection.
- Paper ID: 132.

→ Phylogenetics - Heritability (1 paper): This study found a strong relation between geographic ranges in pairs of closely related mollusc species during the Late Cretaceous, indicate that such trait (geographic range) is heritable at the species level.
- Paper ID: 13.

→ Phylogenetics (2 papers): Both of these papers used phylogenetic methods to assess the origin of multicellularity in volvocine algae. These organisms "diverged relatively recently from unicellular relatives and extant species display a range of intermediate grades between unicellular and multicellular" [59]. Both studies are among the few studies directly using a MLS framework to directly quantify these evolutionary transitions in individuality.
- Papers ID's: 59, 62.

3. Indirect Genetic Effects (IGE) (No. papers = 29; 2008 – 2023)
- Description: An IGE has been defined as the "effect of a gene in the genome of one individual on the phenotype of another individual" (Wade 2026). IGE's sometimes are also deemed as "social

genetic effects” . MLS *in situ* studies under this category were conducted mostly in farm animals (22) in contrast with wild animals (7). A meta-analysis on this subject was recently published by Santostefano et al. (2024). A typical IGE study collects population and trait and/or loci data to assess the effects of interacting partners on a focal individual traits’ and/or reproduction and/or a fitness proxy. The studies in farm animals under this category were all population assessments (i.e., heritability, reproduction, trait variation) and data analyses without any experimental manipulation. In the studies where heritability was estimated, when IGEs where included, its total value always increased (from low values in narrow-sense heritability to high values in total heritable variance; Santostefano et al. 2024).

→ <u>IGE – Wild animals (7 papers)</u>: IGE studies conducted with wild animals’ populations assessments.
- Papers ID’s: 124, 143, 145, 184, 212, 234, 260.
→ <u>IGE – Farm animals (22 papers)</u>: IGE studies conducted with farm animals’ populations assessments.
- Papers ID’s: 57, 85, 93, 104, 107, 115, 159, 169, 176, 178, 188, 189, 193, 198, 202, 207, 208, 213, 218, 229, 257, 268.

4. Group effects (No. papers = 18; 1994 – 2024):
- Description: These studies were all conducted in natural populations of wild animals. In general, these studies assessed the effects of group emergent properties (like networks of interactions or group structure) on focal individuals’ trait variation and/or individual fitness. Thus, and similarly to IGE studies, on these studies, individual trait variation and/or fitness is a function of both the focal individual traits and of the characteristics (composition) of the group to which it belongs. In addition, a subset of these studies directly calculated *Group heritability*, while other subset measured *colony-level selection*, and finally, other subset consisted of *Field experiments*.

Group heritability (No. papers = 2; 2020 – 2021): According to Okasha (2003): “just as individual heritability means the tendency of offspring to resemble their parents with respect to organismic traits, so group heritability should mean the tendency of offspring groups to resemble their parental groups with respect to group-level traits”. Both of these studies measured group heritability.

→ <u>Group heritability – Wild animals (1 paper)</u>: This study found moderate heritability estimates for ants collective behaviours (ranging from 0.17 to 0.32).
- Paper ID: 215.
→ <u>Group heritability – Farm animals (1 paper)</u>: In this study, large-scale feed intake (FI) data was recorded in pigs. The authors found that their results “support the hypothesis that group- and individual-level FI are different traits”.
- Paper ID: 227.

Group effects (No. papers = 9; 2007 – 2024): These studies assessed the effects of group emergent properties (like networks of interactions or group structure) on focal individuals’ trait variation and/or individual fitness.

→ Social networks – Wild animals (4 papers): These studies assessed how social interaction networks affected (or not) focal individuals traits, for example their dispersion capacity [279] and weaponry [226]. Some studies also directly assessed the effects of social networks on individual reproduction [265, 266].
- Papers IDs: 226, 265, 266, 279.

→ Selection – Wild animals (4 papers): These studies directly measured the effects of the social environment ('social selection') on focal individuals' fitness.
- Papers IDs: 55, 123, 158, 275.
→ Group effects – Wild animals (1 paper): This study assessed how "social and ecological factors affecting cold tolerances in range-shifting populations of the female-polymorphic damselfly *Ischnura elegans* in north-east Scotland".
- Paper ID: 197.

Colony selection (No. papers = 3; 2023): These studies directly measured phenotypic variation at the whole-colony level, finding significant trait variation and selection at the colony level.

→ Trait variation (1 paper): This study finds that toxin diversity contained in *Rhytidoponera metallica* and colonies is maintained by interdemic selection that that favours colonies that can exploit more resources and defend against a wider range of predators.
- Paper ID: 267.
→ Selection (1 paper): This study finds shifts in mean trait values along environmental gradients. Such shifts in trait means were mostly coordinated across organizational levels (i.e. worker, colony and species).
- Paper ID: 254.
→ Personality (1 paper): The authors studied the variation in activity levels and exploration among colonies of the thermophilic ant species *Aphaenogaster senilis*, in two undisturbed, contrasting (and close) habitats (grassland and woodland).
- Paper ID: 255.

Field experiment (No. papers = 4; 1994 – 2019)**:** Albeit experimental, these studies were carried out in the field, thus they are under the 'MLS *in situ*' main category. These four studies assessed group effects on focal individuals' phenotypic variation and/or fitness.

→ Group effects – Wild plants (1 paper): In the annual herb *Impatiens capensis*, this study found a complex interaction with neighbourhood characteristics such as relative height of the neighbours and the expression of early plasticity on focal individuals' fitness. Under certain group conditions (certain average height and density), costly plastic elongation responses remain advantageous.
- Paper ID: 77.
→ Group effects – Wild animals (2 papers): Both of these studies, conducted in a spider [64] and a beetle [21], are compatible with standard MLS models.
- Papers ID's: 21, 64.
→ Group effects – Crops (1 paper): The results of this field experiment carried out with wheat, show that "the success of wheat breeding for higher yields over past 100 years in north-western China has

been in part due to unconscious group selection on root traits", which results in smaller, less branched, and deeper roots.
- Paper ID: 201.

5. Contextual analysis (No. papers = 25; 1995 – 2023)
- Description: Contextual analysis follows the methods for analysing phenotypic selection originally developed by Lande and Arnold (1983) and Arnold and Wade (1984), where a multiple regression of relative fitness on phenotype is performed (Goodnight et al. 1992). Contextual analysis extends such methods by including "contextual" or "emergent" traits, that is, traits measured on the group or neighbourhood, in the multiple regression. In this way, relative fitness is a function of individual and group or emergent traits. This phenotypic selection tool allows to disentangle the strength and direction of selection operating at the individual and group levels.

→ Unaltered – Humans (1): So far, this constitutes the only contextual analysis paper done in humans populations, using the Utah Population Database.
- Paper ID: 102.
→ Unaltered (11 papers): Contextual analysis studies done in natural populations, which were not manipulated or altered in any way.
- Papers ID's: 24, 25, 26, 50, 75, 86, 121, 127, 154, 214, 225.
→ Manipulated (13 papers): Contextual analysis studies done in natural populations which were subjected to manipulations, like altering density or spatial structure/organization or resource availability.
- Papers ID's: 37, 49, 51, 52, 56, 68, 70, 74, 80, 82, 134, 231, 252.

MLS Experimental studies

1. Lab experiment (No. papers = 107; 1976 – 2024)
- Description: This is a broad category composed by 24 sub-categories and 84 papers. Some lab experiments imposed group selection and individual selection regimes, comparing responses to selection ('group selection treatments'), the molecular consequences of such treatments, group effects on focal individuals' fitness, microbial culture treatments, and measurements of different aspects of colony-level selection (trait variation, fitness, among others). A subset of these experiments, assessed the effects of Indirect Genetic Effects (IGE) on focal individuals phenotypic variation and/or fitness.

→ Population heritability – Wild animals (1 paper): On this paper, phase III of Wright's Shifting balance process was investigated using meta-populations of the flour beetle, *Tribolium castaneum*. Specifically, popoulation heritability of mean fitness was estimated by the regression of offspring deme means on the weighted parental means. The authors found moderate levels of demic heritability one (0.641-0.690) and two (0.518-0.552) generations after migration.
- Paper ID: 36.
→ Molecular sequencing (lab experiment) – Virus' RNA (1 paper): The authors adapted the Coxsackie virus B3 to a highly permissive and less permissive environment. The authors show that "adaptation occurs by selection of a dominant mutation followed by group selection of minority variants that together, confer the fitness increase observed in the population".
- Paper ID: 126.

→ Molecular sequencing (lab experiment) – Selfish genetic elements (1 paper): The authors addressed the effects of environment on cheater dynamics across different selection levels. In particular, they focused on food availability in *Caenorhabditis elegans*. They found that the proliferation of selfish mitochondrial DNA within hosts depends on nutrient status.
- Paper ID: 206.
→Molecular sequencing (lab experiment) – Colonies' microsatellites (1 paper): This study implemented microsatellites sequencing to determine the genetic identity of somatic and gametic tissues within vascularly fused *Botryllus schlosseri* chimeras, a colonial tunicate. The authors demonstrate that primitive germ cell and somatic cell lineages of this tunicate have traits "that also make them likely units of natural selection".
- Paper ID: 38.
→ Molecular sequencing (lab experiment) – Alleles (1 paper): The authors "isolated 109 *Enterococcus faecalis* from chicken faecal samples in 6 provinces of China to investigate the prevalence and transmission mechanism of the bacitracin resistance locus bcrABDR in *E. faecalis*.". The results indicated "that multiple levels of selection regulate the spread of bcrABDR among *E. faecalis* populations".
-Paper ID: 141.
→ Microbiome assessment – Wild animals (1 paper): The authors tested whether aggregation behaviours may be linked to the gut microbiota of the host. The authors found that "a host's gut microbiota is not necessarily a major driver or a consequence of aggregation behaviour in species with inter-individual variation in group living".
- Paper ID: 272.
→ Group selection treatments – Wild plants (2 papers): These studies imposed individual and group selection regimes (lines) through several generations. In both cases, the response to group selection was always in the expected direction.
- Papers ID's: 11, 146.
→ Group selection treatments – Wild animals (18 papers): Historically, these studies constitute the first empirical evidence of multilevel selection in the lab [1 – 10] (these ten references span from 1976 to 1984). Fifteen of these 18 studies were carried out with beetles. Similarly as group selection treatments imposed in plants (see above), on these studies, individual and group selection regimes were artificially imposed on lab populations. While the strength of selection varies across studies, overall, imposed group selection was usually effective and had stronger phenotypic effects than individual selection.
- Papers ID's: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20, 40, 46, 60, 149.
→ Group selection treatments – Virus (1 paper): This study applied individual and group selective pressures, finding evidence that group selection modulates the virulence of vesicular stomatitis virus populations, "in opposition to an unlimited increase in virulence by competitive optimization promoted by individual selection".
- Paper ID: 35.
→ Group personality – Wild animals (1 paper): This study investigated whether individual and group level morphology and behaviour traits (e.g., size of the largest individual, average activity behavioural type in the group) could partially explain the variability in how rapidly groups transition through different mating systems (i.e., small groups often transition into a system in which a single, large male monopolizes mating opportunities). The authors found that groups with males that exhibit higher social plasticity tended to take longer to transition to a mating system dominated by a single male. They conclude that "groups in identical ecological conditions can diverge in their mating

systems based on how much individuals in the group change their behaviour in response to the behaviour of other members of the group".
- Paper ID: 192.
→ <u>Group fitness – Wild animals (1 paper)</u>: This study assessed the effects of worker density on group direct benefits (group reproduction and survival) in the termite species *Reticulitermes chinensis*. The authors investigated group benefits regarding worker density (11 different densities), and measured the group benefits and resource consumption with the same group members in two types of artificial nest areas. The results suggest that "group benefits influenced by the high worker density may at least partially explain the group living of eusocial insects in ecology".
- Paper ID: 220.
→ <u>Group effects (Social networks) – Wild animals (6 papers)</u>: These studies experimentally manipulated the social networks of captive populations, measuring the trait and/or fitness. Alternatively, some studies implemented network analyses of captive populations to assess the effects of network topography and structure on focal individuals' reproduction. Other studies manipulated the age structure of the captive populations and found that fitness is associated with variance in age at three different levels of organization: the individual, interacting social partners, and the population. Overall, these studies show effects of the social environment and the topography of the interaction networks (an emergent property) on individual trait variation and fitness.
- Papers ID's: 78, 79, 155, 190, 251, 273.
→ <u>Group effects – Wild animals (14 papers)</u>: These experimental studies examined how different imposed group properties (i.e., group size, social interactions, group collective behaviour *aka* personality, phenotypic composition, among others) affects both individual and group-level traits (for example the group's mating system or overall group activity, among others). Several of these studies were conducted in water striders.
- Papers ID's: 22, 54, 96, 114, 122, 135, 157, 160, 161, 172, 211, 223, 232, 233.
→ <u>Group effects – Farm animals (3 papers)</u>: These studies with farm animals tested the effects of imposed group properties (i.e., social interactions, group's aggressiveness/docility, groups' fearfulness) affects both individual and group-level traits.
- Papers ID's: 84, 113, 277.
→ <u>Group effects – Crops (2 papers)</u>: Both of these studies were conducted in China, with wheat crop-land. These studies show how, depending on deme density or interaction with neighbourhood roots, individual plants change root allocation. Furthermore, these studies show that older wheat varieties have been subjected to "inadvertent group selection" [200].
- Papers ID's: 200, 249.
→ <u>Culture experiment – Fungi (1 paper)</u>: In the fungus *Neurospora tetrasperma*, selection can act at different levels: nuclei compete in their replication and transmission into short-lived asexual spores, but, at the level of the heterokaryotic individual, nuclear types cooperate to produce long-lived sexual spores. The authors evaluated the extent of MLS in three *N. tetrasperma* strains, by assessing the ratio between nuclear types (homokaryons: heterokaryons) under different conditions, measuring fitness traits of homo- and heterokaryotic mycelia with varying nuclear ratios. They found that the two nuclei have complementary traits, reflecting labour division and cooperation.
- Paper ID: 177.
→ <u>Culture experiments – Bacteria (5 papers)</u>: In bacterial cultures, these experiments manipulated resource availability, while measuring competition and/or cooperation dynamics. These studies overall show that bacterial populations develop mechanisms to avoid 'cheaters' overtake. The tools employed include the Price Equation [90] or analysis of quorum sensing [138], among others.

- Papers ID’s: 88, 90, 138, 196, 238.

→ Culture experiment – Algae (1 paper): In this study, it was tested the hypotheses that exotosins, in algae, are either exploitable public goods that benefit all cells, regardless of investment, or non-exploitable private goods involved in cell-level functions. The authors subjected two mixotrophic microalga *Prymnesium parvum* strains, contrasting in their toxicity, to two competitors: green algae and diatoms. They show that *P. parvum* strains with high toxicity actually benefit from dense populations of competing green algae, which can also be prey for *P. parvum*, yielding a relative fitness advantage over coexisting strains with lower toxicity. The authors conclude that “An evolutionary unstable, ecologically devastating public good may emerge from traits selected at lower levels expressed in novel environments”.

- Paper ID: 99.

→ Community selection treatments – Wild animals (2 papers): These two pioneer studies were conducted with simple communities of two beetle species (*Tribolium castaneum* and *T. confusum*). In these experiments, a series of one-generation assays were performed on these co-evolved simple communities. Selection at the community level is performed on one trait measured in one species and correlated responses in other traits measured both within species and among species are monitored. The studies demonstrate community-level selection, defined as the differential survival and/or reproduction of entire communities, resulting in changes in the phenotype of a community.

- Papers ID’s: 18, 19.

→ Community heritability – Wild animals (1 paper): This study, also used laboratory co-existing populations of two beetle species (*Tribolium castaneum* and *T. confusum*), with different treatments for a competitive phase study: mixed-species communities (n=10), single-species populations (n=10), or in the standard culture conditions used to maintain stocks in the laboratory. Each community was replicated 15 times, in order to provide an assessment of the distribution of the competitive outcome. Lineages within the treatments became highly differentiated for all measures of competitive outcome. The fraction of the variance among lineages constitutes the community heritability. This study demonstrated high levels of community heritability in the measures of competitive outcomes.

- Paper ID: 29.

→ Colony trait variation – Wild animals (8 papers): The eight studies investigated variation in colony traits: queen transport, termite next architecture, maintenance, personality, colony size, aggregate aggressive behaviour, among others. Colony-level traits were related to reproduction both at the individual and colony levels.

- Papers ID’s: 41, 101, 129, 139, 163, 191, 271, 278.

→ Colony selection – Wild animals (8 papers): In line with the previous sub-category, these studies show natural selection occurring at the colony level, for example by showing how “colony-level personality may be linked with differences in colony life-history strategy, which may promote the emergence and maintenance of personality traits in group-living species” [140]. Or by showing how individual phenotypes are properties of the composite "sociogenome" [63]. Other studies showed how colony personality influenced colony productivity [230] or fitness [81]. Overall, these studies show that natural selection occurs simultaneously at two levels: the individual insect and the colony levels, and how trait variation at both levels influences reproductive outcomes.

- Papers ID’s: 63, 81, 87, 98, 140, 230, 245, 258.

→ Colony fitness – Wild animals (2 papers): Both of these studies explicitly show how the fitness of the whole colony is significantly affected by the proportion of worker types (subordinate, dominant) [16] or by population structure [270].

- Papers ID’s: 16, 270.

→ Clonal lineages – Fungi (1 paper): This study showed that natural selection occurs at the level of individual nuclei in the symbiont fungi *Glomus etunicatum* (Phylum Glomeromycota), by investigating its nuclear dynamics using live three-dimensional imaging and mathematical models. Spores of this fungi are populated by an influx of a stream of nuclei from the surrounding mycelium rather than by divisions of a single founder nucleus.
- Paper ID: 71.
→ Clonal lineages – Algae (1 paper): This study used the red macro-algae *Delisea pulchra* and *Asparagopsis armata* to test whether modules of these algae have replication, heritable variation, and that selective agents distinguish among the variants (aka being a unit of selection). The authors detected significant phenotypic variation among modules for fitness-related traits (growth, secondary metabolite concentrations, and rates of tissue loss to herbivory) in each species and significant heritability estimates for secondary metabolite production and tissue loss to herbivory in *D. pulchra*. This study shows that selection acts on phenotypic variation among algae modules within individuals, affecting evolutionary change.
- Paper ID: 53.

IGE experiment (No. papers = 23; 1987 – 2023): Controlled experiments done to assess the effects of IGEs on focal individuals phenotypic variation and/or fitness.

→ IGE experiment – Wild animals (14 papers): controlled IGE experiments done with wild animals.
- Papers ID's: 14, 94, 108, 110, 117, 131, 144, 164, 168, 181, 210, 243, 253, 269.
→ IGE experiment – Plants (1 paper): controlled IGE experiment done with plants.
- Paper ID: 261.
→ IGE experiment – Microbes (1 paper): controlled IGE experiment done with microorganisms.
- Paper ID: 69.
→ IGE experiment – Lab mice (1 paper): controlled IGE experiment done with lab mice.
- Paper ID: 221.
→ IGE experiment – Farm animals (6 papers): controlled IGE experiments done with farm animals.
- Papers ID's: 106, 111, 120, 128, 133, 162.

2. Breeding through group selection (No. papers = 32; 1996 – 2023)
- Description: Typical studies under this category had two contrasting breeding treatments: individual-based breeding (classical way to breed animals) and group-based breeding. Classically, animals are breed by selecting their desired individual traits (individual-based breeding) through several generations. However, these studies also had lines selected desired group traits such as group productivity, reduced social stress, low mortality, among others. These imposed group selection regimes are always successful, and some of the studies report that they cause genetic and physiological changes on focal individuals – in addition to affecting their reproductive success. Twenty of these 21 papers were conducted in farm animals (chickens and pigs), with just one study [103], conducted in the Japanese quail. A subset of these studies consisted of breeding programs that calculated IGE
- Papers ID's: 27, 28, 30, 31, 32, 33, 44, 45, 48, 61, 66, 73, 100, 103, 112, 119, 156, 175, 219, 224, 262.

IGE breeding (No. papers = 11; 2008 – 2019): these studies consisted on breeding programs that calculated IGEs, taking them into account to breed the next generations of animals or crops.

When estimations of heritable variance include both direct and indirect genetic effects, the values were always higher than when only direct effects are considered. Selection based or considering IGEs seem to target several traits simultaneously, rather than a single trait [95].

→ IGE breeding – Wild animals (1 paper): breeding program that calculated IGEs for wild animals (grasshoppers).
- Paper ID: 182.
→ IGE breeding – Farm animals (9 papers): breeding program that calculated IGEs for farm animals.
- Papers ID's: 58, 83, 89, 95, 109, 136, 137, 152, 170.
→ IGE breeding – Crops (1 paper): breeding program that calculated IGEs for tree plantations.
- Paper ID: 97.

3. Psychology experiment (No. papers = 19; 2011 – 2024)

- Description: these psychological experiments followed the cultural multilevel selection framework (Wilson et al. 2023). Some studies explicitly made player groups compete with each other [180], while others used the cultural MLS framework to design the REDD+ forest restoration program in Tanzania [167]. In general these studies implemented games where, either, groups of players competed, or the effects of group properties (i.e., size) on individual-level cooperative behaviour were quantified.
- Papers ID's: 76, 118, 165, 166, 167, 171, 180, 185, 186, 199, 216, 222, 228, 236, 239, 240, 244, 247, 276.

4. Artificial selection (No. papers = 22; 2000 – 2023)

- Description: in artificial selection studies, humans breed whole communities (like microbiomes) or species consortia or aggregates (like yeast aggregates) for specific desired traits (like bigger colony size, for yeasts), under specific environmental conditions.

→ Single species (4 papers): artificial selection studies done with one species.
- Papers ID's: 65, 92, 150, 151.
→ Multicellularity (12 papers): these studies involve applying selection pressures to microbe populations to study adaptation. In particular, certain environmental pressures (for example, low/no oxygen or increased gravitational forces, or predators pressures) are imposed in test tube populations, composed by single-cell microbes. Usually, such microbes at first increase in size, and then start to aggregate, forming very tight colonies or aggregates, that ultimately have transition towards a multicellular state, even showing labour division. Nine of these 12 studies were done with yeasts.
- Papers ID's: 91, 105, 116, 130, 153, 173, 179, 187, 195, 204, 250, 263.
→ Microbiome (5 papers): artificial selection studies done with microbiomes composed by hundreds or thousands of species belonging to different taxa and functional guilds.
- Papers ID's: 42, 43, 125, 194, 242.
→ Consortia (1 paper): artificial selection studies done with a consortia of species.
- Paper ID: 205.

Supplementary Table S3. Summary of 26 contextual analysis of phenotypic selection done between 1995 and 2025. Beta regression coefficients show the effects of organismal ‘size’ (or traits which are a proxy of size, like height) at the individual (“Ind. β Coef.”) and group (“Group β Coef.”) levels, on individual fitness (or fitness proxies). Lower and higher estimates of the regression coefficients are also shown. Several studies measured additional traits (“Other traits”), in which the direction of natural selection (i.e., the sign of the β coefficient) is indicated (+ or - signs) at the individual (I) and group (G) levels.

Reference	**Ind. β Coef.**	**Ind. Low**	**Ind. High**	**Group β Coef.**	**Group Low**	**Group High**	**Other traits**
Stevens et al. (1995) Am Nat	1.130	0.510	1.740	-1.987	-3.030	-0.520	
Tsuji (1995) Am Nat	0.071	0.071	0.071	-0.110	-0.110	-0.110	
Banschbach and Herbers (1996) Evolution	Estimates not reported						
Herbers and Banschbach (1999) Behav Ecol Sociobiol	Estimates not reported						
Volis et al. (2002) Biol J Linn Soc	0.320	0.320	0.320	-6.420	-6.420	-6.420	Size of reproductive organs (I+,_G-)
Aspi et al. (2003) Evolution	0.695	0.589	0.800	0.329	0.187	0.471	Stem number(I+,G-)
Donohue (2003) Am Nat	2.140	2.140	2.140	2.150	2.150	2.150	
Donohue (2004) Ecology	-1.520	-1.520	-1.520	2.110	2.110	2.110	Stem mass (I+,G-)
Weinig et al. (2007) Evolution	0.610	0.400	0.820	-0.585	-0.750	-0.420	Elongation (I+,G-), lifespan (I+,G+), basal branching (I+,G-)
Boege (2010) Ecology	0.900	0.900	0.900	-0.750	-0.750	-0.750	Herbivory resistance (I+,G-)
Eldakar et al. (2010) Evolution	0.752	0.000	0.000	-0.556	0.000	0.000	Male aggression (I+,G-)
Eldakar and Gallup (2011) PLoS One	0.393	0.000	0.000	-0.485	0.000	0.000	Male aggression (I+,G-)
Formica et al. (2011) Evolution	0.378	0.378	0.378	-0.356	-0.356	-0.356	
Stinson et al. (2011) Ecosphere	0.350	0.350	0.350	0.530	0.530	0.530	
Bolstad et al. (2012) Ecol Evol	2.540	-0.110	3.570	-1.080	-1.290	1.750	Growth rate (I+,G+)
Laiolo and Obeso (2012) PloS One	-0.600	0.000	0.000	0.580	0.000	0.000	Song repertoire size (I-,G+)
Moorad (2013) Evolution	Size or a proxy of size were not evaluated						Polygyny (I+,G+), Polyandry (I+,G+)
Searcy et al. (2014) Ecology	Size or a proxy of size were not evaluated						
Bouwhuis et al. (2015) J Anim Ecol	0.069	0.069	0.069	0.176	0.176	0.176	

Weis et al. (2015) Evolution	Size or a proxy of size were not evaluated						
Fisher et al. (2017) Evolution	0.239	0.000	0.000	0.105	0.000	0.000	Growth rate (I+,G+)
Santostefano et al. (2020) Evolution	0.530	0.530	0.530	-0.570	-0.570	-0.570	Docility (I-,G-)
Fisher et al. (2021) Proc R Soc B	0.239	0.000	0.000	0.105	0.000	0.000	
Marín (2021) Genet Resour Crop Evol	0.311	0.224	0.453	0.707	0.410	1.005	1st leaf angle (I+,G-)
Costello et al. (2023) Evolution	0.134	0.011	0.351	0.028	0.001	0.054	
Philson et al. (2025) Proc R Soc B	-0.182	-0.942	0.578	-0.263	-1.648	0.908	Closeness (I+, G-)

References

- Arnold SJ, Wade MJ. 1984 On the measurement of natural and sexual selection: theory. *Evolution* **38**, 709-719. doi:10.1111/j.1558-5646.1984.tb00344.x
- Aspi J, Jäkäläniemi A, Tuomi J, Siikamäki P. 2003 Multilevel phenotypic selection on morphological characters in a metapopulation of *Silene tatarica*. *Evolution* **57**, 509-517. doi:10.1111/j.0014-3820.2003.tb01542.x
- Banschbach VS, Herbers JM. 1996 Complex colony structure in social insects: II. Reproduction, queen-worker conflict, and levels of selection. *Evolution* **50**, 298-307. doi:10.1111/j.1558-5646.1996.tb04493.x
- Boege K. 2010 Induced responses to competition and herbivory: natural selection on multi-trait phenotypic plasticity. *Ecology* **91**, 2628-2637. doi:10.1890/09-0543.1
- Bolstad GH, Pelabon C, Larsen LK, Fleming IA, Viken Å, Rosenqvist G. 2012 The effect of purging on sexually selected traits through antagonistic pleiotropy with survival. *Ecol. Evol.* **2**, 1181-1194. doi:10.1002/ece3.246
- Bouwhuis S, Vedder O, Garroway CJ, Sheldon BC. 2015 Ecological causes of multilevel covariance between size and first-year survival in a wild bird population. *J. Anim. Ecol*. **84**, 208-218. doi:10.1111/1365-2656.12264
- Costello RA, Cook PA, Brodie III ED, Formica VA. 2023 Multilevel selection on social network traits differs between sexes in experimental populations of forked fungus beetles. *Evolution* **77**, 289-303. doi:10.1093/evolut/qpac012
- Damuth J, Heisler IL. 1988 Alternative formulations of multilevel selection. *Biol. Philos.* **3**, 407-430. doi:10.1007/BF00647962
- Donohue K. 2003 The influence of neighbor relatedness on multilevel selection in the Great Lakes sea rocket. *Am. Nat.* **162**, 77-92. doi:10.1086/375299
- Donohue K. 2004 Density-dependent multilevel selection in the great lakes sea rocket. *Ecology* **85**, 180-191. doi:10.1890/02-0767
- Eldakar OT, Gallup AC. 2011 The group-level consequences of sexual conflict in multigroup populations. *PloS One* **6**, e26451. doi:10.1371/journal.pone.0026451
- Eldakar OT, Wilson DS, Dlugos MJ, Pepper JW. 2010 The role of multilevel selection in the evolution of sexual conflict in the water strider *Aquarius remigis*. *Evolution* **64**, 3183-3189. doi:10.1111/j.1558-5646.2010.01087.x
- Fisher DN, Boutin S, Dantzer B, Humphries MM, Lane JE, McAdam AG. 2017 Multilevel and sex-specific selection on competitive traits in North American red squirrels. *Evolution* **71**, 1841-1854. doi:10.1111/evo.13270
- Fisher DN, LeGrice RJ, Painting CJ. 2021 Social selection is density dependent but makes little contribution to total selection in New Zealand giraffe weevils. *Proc. R. Soc. B: Biol. Sci.* **288**, 20210696. doi:10.1098/rspb.2021.0696
- Formica VA, McGlothlin JW, Wood CW, Augat ME, Butterfield RE, Barnard ME, Brodie III ED. 2011 Phenotypic assortment mediates the effect of social selection in a wild beetle population. *Evolution* **65,** 2771-2781. doi:10.1111/j.1558-5646.2011.01340.x
- Goodnight CJ, Schwartz JM, Stevens L. 1992 Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *Am. Nat.* **140**, 743-761. doi:10.1086/285438

- Herbers JM, Banschbach VS. 1999 Plasticity of social organization in a forest ant species. *Behav. Ecol. Sociobiol.* **45**, 451-465. doi:10.1007/s002650050584
- Laiolo P, Obeso JR. 2012 Multilevel selection and neighbourhood effects from individual to metapopulation in a wild passerine. *PLoS One* **7**, e38526. doi:10.1371/journal.pone.0038526
- Lande R, Arnold SJ. 1983 The measurement of selection on correlated characters. *Evolution* **37**, 1210-1226.
- Marín C. 2021 Spatial and density-dependent multilevel selection on weed-infested maize. *Genet. Resourc. Crop Evol.* **68**, 885-897. doi:10.1007/s10722-020-01031-1
- Marín C. 2024 Three types of units of selection. *Evolution* **78**, 579-586. doi:10.1093/evolut/qpad234
- Montazeri A, Mohammadi S, M.Hesari P, Ghaemi M, Riazi H, Sheikhi-Mobarakeh Z. 2023 Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. *Syst. Rev.* **12**, 239. doi:10.1186/s13643-023-02410-2
- Moorad JA. 2013 Multi-level sexual selection: individual and family-level selection for mating success in a historical human population. *Evolution* **67**, 1635-1648. doi:10.1111/evo.12050
- Okasha S. 2003 The concept of group heritability. *Biol. Philos.* **18**, 445-461. doi:10.1023/A:1024140123391
- Okasha S. 2006 *Evolution and the levels of selection*. New York, United States: Oxford University Press.
- Philson CS, Martin JG, Blumstein DT. 2025 Multilevel selection on individual and group social behaviour in the wild. *Proc. R. Soc. B: Biol. Sci.* **292**, 20243061. doi:10.1098/rspb.2024.3061
- Price GR. 1970 Selection and covariance. *Nature* **227**, 520-521. doi:10.1038/227520a0
- Santostefano F, Garant D, Bergeron P, Montiglio PO, Reale D. 2020 Social selection acts on behavior and body mass but does not contribute to the total selection differential in eastern chipmunks. *Evolution* **74**, 89-102. doi:10.1111/evo.13875
- Santostefano F, Moiron M, Sánchez-Tójar A, Fisher DN. 2025 Indirect genetic effects increase the heritable variation available to selection and are largest for behaviors: a meta-analysis. *Evol. Lett.* **9**, 89-104. doi:10.1093/evlett/qrae051
- Searcy CA, Gray LN, Trenham PC, Shaffer HB. 2014 Delayed life history effects, multilevel selection, and evolutionary trade-offs in the California tiger salamander. *Ecology* **95**, 68-77. doi:10.1890/13-0120.1
- Stevens L, Goodnight CJ, Kalisz S. 1995 Multilevel selection in natural populations of *Impatiens capensis*. *Am. Nat.* **145**, 513-526. doi:10.1086/285753
- Stinson KA, Brophy C, Connolly J. 2011 Catching up on global change: new ragweed genotypes emerge in elevated CO2 conditions. *Ecosphere* **2**, 1-11. doi:10.1890/ES10-00168.1
- Tsuji K. 1995 Reproductive conflicts and levels of selection in the ant *Pristomyrmex pungens*: contextual analysis and partitioning of covariance. *Am. Nat.* **146**, 586-607. doi:10.1086/285816
- Volis S, Mendlinger S, Ward D. 2002 Differentiation along a gradient of environmental productivity and predictability in populations of *Hordeum spontaneum* Koch: multilevel selection analysis. *Biol. J. Linn. Soc.* **75**, 313-318. doi:10.1046/j.1095-8312.2002.00021.x
- Wade MJ. 2026 Sewall Wright and the Shifting Balance Theory. In: Wolf JB, De Moraes Russo CA (Eds). *Encyclopedia of Evolutionary Biology (Second Edition)*. Amsterdam, the Netherlands: Academic Press. Pp. 387-395. doi:10.1016/B978-0-443-15750-9.00111-7

- Weinig C, Johnston JA, Willis CG, Maloof JN. 2007 Antagonistic multilevel selection on size and architecture in variable density settings. *Evolution* **61**, 58-67. doi:10.1111/j.1558-5646.2007.00005.x
- Weis AE, Turner KM, Petro B, Austen EJ, Wadgymar SM. 2015 Hard and soft selection on phenology through seasonal shifts in the general and social environments: a study on plant emergence time. *Evolution* **69**, 1361-1374. doi:10.1111/evo.12677
- Wilson DS, Madhavan G, Gelfand MJ, Hayes SC, Atkins PW, Colwell RR. 2023 Multilevel cultural evolution: From new theory to practical applications. *Proc. Natl. Acad. Sci. USA* **120**, e2218222120. doi:10.1073/pnas.2218222120
- Wilson DS, Wilson EO. 2007 Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* **82**, 327-348. doi:10.1086/522809