What do isogamous organisms teach us about sex and the two sexes?

Jussi Lehtonen¹, Hanna Kokko² and Geoff A. Parker³

¹Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia
²Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
³Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

Isogamy is a reproductive system where all gametes are morphologically similar, especially in terms of size. Its importance goes beyond specific cases: to this day non-anisogamous systems are common outside of multicellular animals and plants, they can be found in all eukaryotic super-groups, and anisogamous organisms appear to have isogamous ancestors. Furthermore, because maleness is synonymous with the production of small gametes, an explanation for the initial origin of males and females is synonymous with understanding the transition from isogamy to anisogamy. As we show here, this transition may also be crucial for understanding why sex itself remains common even in taxa with high costs of male production (the twofold cost of sex). The transition to anisogamy implies the origin of male and female sexes, kickstarts the subsequent evolution of sex roles, and has a major impact on the costliness of sexual reproduction. Finally, we combine some of the consequences of isogamy and anisogamy in a thought experiment on the maintenance of sexual reproduction. We ask what happens if there is a less than twofold benefit to sex (not an unlikely scenario as large short-term benefits have proved difficult to find), and argue that this could lead to a situation where lineages that evolve anisogamy—and thus the highest costs of sex—end up being associated with constraints that make invasion by asexual reproduction unlikely (the ‘anisogamy gateway’ hypothesis).

This article is part of the themed issue ‘Weird sex: the underappreciated diversity of sexual reproduction’.

1. Introduction

Isogamy (see box 1 for glossary of definitions of terms used) is a reproductive system where all gametes are morphologically similar, particularly in terms of size, and there is no separation into male and female gametes. It may quite reasonably seem like an ‘unusual’ reproductive system, given its apparent rarity. Yet we, along with all other anisogamous organisms, are almost certainly derived from isogamous ancestors (e.g. [1–4]), and isogamy is still common today in unicellular organisms (e.g. [5]).

Nevertheless, we live in a world where sexual dimorphism of some kind is seemingly the norm. Even when there are no separate male and female sexes as such (i.e. simultaneous hermaphrodites), there is dimorphism at the gamete level, where clearly diverged male and female gametes are produced—male gametes being by definition the smaller ones. The majority of research on topics such as sexual selection, sexual conflict, the evolution and maintenance of sexual reproduction, parental care and many others assumes, either explicitly or implicitly, diverged male and female gametes.

While anisogamy is almost universal in complex multicellular eukaryotes (with some notable exceptions in multicellular algae, e.g. [6]), the opposite is true when we move into the world of unicellular organisms. Here, many of the asymmetries that are prevalent in multicellular organisms disappear, including the asymmetry in gametes. Most eukaryote lineages are microbial...
**Box 1. Glossary of terms and definitions used.**

**Anisogamy:** Size dimorphism of gametes: one gamete type is larger (e.g. ova) than the other (e.g. spermatozoa), and gametic fusion (now) occurs only between the larger and the smaller gametes.

**Facultative sex:** Sex that is not obligate; facultatively sexual organisms can perform both asexual and sexual life cycles.

**Hermaphroditism:** (in animals). **Monogy** (in plants): Systems in which male and female functions co-occur in one individual, i.e. a single individual produces both small and large gametes.

**Heterothallism:** Reproductive systems in fungi where syngamy can only occur between haploid cells carrying different mating-type alleles.

**Homothallism:** A reproductive system in fungi that is often simply called self-compatibility; by a stricter definition, a homothallic, haploid individual is able to mate with its own mitotic descendants [16].

**Isogamy:** All gametes in isogamous gametic systems are of similar size. They are not identical, however, as isogamy is almost always associated with mating types. The word originates from iso = equal + gamia (from the Greek games) = marriage.

**Karyogamy:** The fusion of two gametic nuclei.

**Mating types:** Mating types are gamete genotypes for molecular mechanisms that regulate compatibility between fusing gametes. Mating types guarantee disassortative fusion in both isogamous and anisogamous gametic systems: if the gametes are composed of + and − mating types, then only fusions of + and − gametes are possible. Isogamous species often have two mating types, but some species have several mating types, even up to hundreds.

**Separate sexes, gynodioecy (in animals), dioecy (in plants):** Systems in which the two sexes (males and females) are separate, and male individuals by definition produce small gametes and female individuals produce large gametes.

**Syngamy:** The fusion of two gametes to form a zygote.

unicells [7,8], and here isogamy is the norm, meaning that all gametes (i.e. cells that fuse to form a zygote) are of similar size. Furthermore, and very importantly from an evolutionary and ecological perspective, this implies that in most such cases the parental investment from the two parents is equal, unlike in anisogamous systems where the female gamete usually provides the majority of non-genetic resources required for initial zygote development.

Here, we review isogamy, some of its essential features, outline where it can be found, and the many fascinating ways in which life in the isogamous world differs from our ‘norm’.

Our story begins with the last common ancestor of all eukaryotes. All eukaryotes either reproduce sexually or are descended from a sexually reproducing ancestor [9–15]. Sex can involve two processes that are particularly relevant here: in syngamy, gametes fuse to form a zygote and in karyogamy, two gametic nuclei fuse. Both processes occur in virtually all higher animals and plants, but in some organisms (e.g. some ciliates, some fungi) sex involves karyogamy but not syngamy. Because this article is largely focused on how isogamy contrasts with the ‘normal’ world (anisogamy), we focus on eukaryotes with both syngamy and karyogamy. Such cases can be typically categorized as either having only one isogamous sex (defined as just one size of gamete) but with two or more mating types (see box 1 and next section), or as having two sexes as a result of anisogamy, where males produce smaller and females larger gametes. The two sexes can coexist in one individual, as in simultaneous hermaphrodites (or male and female ‘functions’; [17,18]). Most unicellular organisms that produce gametes appear to be isogamous while most multicellular organisms are anisogamous.

Why should we care about isogamy? Why has it persisted against the seemingly ubiquitous rise of anisogamy? We all have sexually reproducing, isogamous ancestors and non-anisogamous systems are very common, though not in multicellular animals and plants. Isogamy has the interesting feature that, assuming no parental care beyond the initial provisioning of the zygote, the latter type of provisioning (i.e. investment in the zygote) is shared equally between the two parents. Anisogamy disrupts this symmetry, creating potential for parental sexual dimorphism and different selective forces on the two sexes, with profound evolutionary and ecological consequences [19,20]. To understand the initial origin of males and females is synonymous with understanding the transition from isogamy to anisogamy—and to understand the transition, we need to understand isogamy.

**2. The diversity of isogamous reproduction: where and in what forms?**

Anisogamy is universal in most sexually reproducing multicellular plants and all multicellular animals. Plants (defined here as Viridiplantae: green algae and land plants) present some exceptions to this: for example, the multicellular colonial green algae genera *Pandorina, Volvox* and *Yamagishiella* are isogamous [21,22] as are some multicellular marine green algae [23], but anisogamy is nevertheless understandably much more familiar to us than isogamy. None of this of course implies that isogamy is rare—but how common is it? An exact answer to this question is not available, given that for a great number of microbial species we do not even know if they reproduce sexually [9]. Despite a strong impression that isogamy prevails in unicellular species, there are many exceptions to this too: e.g. unicellular Bryopsidae green algae are isogamous [23], and as we have seen with the plant examples above, multicellularity does not automatically mean abandoning isogamy. From a broad-scale phylogenetic perspective, all eukaryote super-groups contain isogamous species (table 1).

Finally, a more detailed answer to this question depends on how strictly we define isogamy. A great variety of reproductive systems and life cycles exist in nature, even more so when we look into the isogamous world. It is not surprising then that some definitions in biology, including that for isogamy, do not unambiguously fit all possible reproductive
Table 1. Examples of isogamous organisms in eukaryote super-groups and largest assemblages. Eukaryote classification after Adl et al. [24] and Burki [25], largest assemblages mainly after Burki [25]. To our knowledge, there are no isogamous metazoans, red algae or land plants; otherwise isogamy can be found in all groups presented here. For a summary of the distribution of isogamy and anisogamy in algae, see Bell [6].

<table>
<thead>
<tr>
<th>super-groups</th>
<th>largest assemblages</th>
<th>example of isogamous species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eukaryota</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoebozoa</td>
<td>Fungi</td>
<td>the social amoeba Dictyostelium discoideum [26]</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>Metazoa</td>
<td>baker’s/brewer’s yeast Saccharomyces cerevisiae [27,28]</td>
</tr>
<tr>
<td>Excavata</td>
<td></td>
<td>the parasitic protist (and cause of sleeping sickness) Trypanosoma brucei [29,30]</td>
</tr>
<tr>
<td>Sar</td>
<td>Alveolates</td>
<td>the dinoflagellate Polykrikos kofoidii [31]</td>
</tr>
<tr>
<td>Stramenopiles</td>
<td>Rhizaria</td>
<td>the brown alga Assosira mirabilis [32]</td>
</tr>
<tr>
<td>Archaeplastida</td>
<td>Rhodophyceae (red algae)</td>
<td>the foraminiferan Nummulites venosus [33]</td>
</tr>
<tr>
<td>Viridiplantae</td>
<td>green algae</td>
<td>the green alga Carteria palmata [34]</td>
</tr>
<tr>
<td></td>
<td>land plants</td>
<td>no isogamous species</td>
</tr>
</tbody>
</table>

strategies that exist. In this article, we will mostly focus on isogamy in the sense that it complements anisogamy: that is, gametes are produced in both, while they are size dimorphic in anisogamy, they are similar in size in isogamy. Despite this specific focus, it is nevertheless useful to consider briefly how things can differ from the ‘norm’ under more general settings, including both gametic and non-gametic reproductive systems.

While many familiar sexually reproducing organisms form gametes in the gonads, this is clearly not the case for isogamous unicells, as gonads are specialized glands requiring multicellular structures. For example, in the unicellular algae Chlamydomonas reinhardtii and Carteria palmata, vegetative cells (i.e. the adults themselves) come in two mating types, + and −, and they split into gametes of the corresponding type (e.g. [34–36]). These species are classic examples of unambiguous isogamous reproduction. Both species belong to the order Volvocales, which includes a range of species from unicellular, isogamous to multicellular, strongly anisogamous ones, as well as various degrees of germline/soma differentiation (e.g. [6,37–43]).

Regardless of how gametes in isogamous organisms are produced, they then meet and fuse (syngamy), and undergo nuclear fusion (karyogamy). Fusion necessitates reduction division to restore the original ploidy level from one zygote to the next, and this can occur either before or after syngamy, depending upon whether the organism spends most of its life in the diploid or the haploid stage.

Intermediate, slightly anisogamous forms can also have fascinating and surprising features. For example, the brown alga Ectocarpus (which has emerged as a model organism for the brown algae [44]) produces male and female gametes that differ only very slightly in size, but more markedly in their ‘behaviour’: female gametes rapidly settle and produce a pheromone, while male gametes swim for longer and are attracted to the pheromone [45]. Perhaps even more fascinating is that gametes that fail to fuse can develop parthenogenetically, and in some Ectocarpus species both male and female gametes have this ability [45]. Presumably, this is possible due to the very small difference in gamete size, where female gametes have not yet become the sole resource provider. In some algae, the size difference between male and female gametes relates to the number of cell divisions of the mother cell, which may differ by only one division [46,47].

With many other species, gametes in a strict sense do not exist at all, and it becomes ambiguous whether the term ‘isogamy’ should be used. For example, ciliates have sex by means of exchanging haploid, ‘germline’ micronuclei during conjugation (e.g. [4,9,48]), which means that there is karyogamy but not syngamy. Should this be called isogamous or anisogamous sex, or neither? The migratory micronuclei exchanged between individuals are usually the same size as the retained, stationary micronuclei. In the absence of a size difference, this is not anisogamy. Because the meiotic products are similar in size, the system could be (and often is) considered isogamous under a broad definition, even though gametes are not produced. Despite lack of size dimorphism, terminology involving ‘males’ and ‘females’ has also been used in this setting (e.g. [4,49]): maleness is then associated with the migratory micronucleus, which is transferred to a different individual somewhat analogously to male gametes. Although clearly reminiscent of the travels of pollen and/or sperm, note that this broad usage of the word ‘male’ takes a very different trait, the motility of the vehicle carrying the genetic material, as the basis of maleness, rather than the size-based criterion that is the convention when discussing anisogamy.

Similar considerations apply to, e.g. basidiomycete fungi, where sex can take place between two mycelia, which is again often termed isogamy despite the lack of gametes (e.g. [50]). If we follow this route, and essentially define isogamy as existing whenever one cannot distinguish between small and large gametes (i.e. either because gametes do not differ in size or because there are no gametes), then isogamy may well be a more common reproductive strategy than the more familiar anisogamy. As always, it is important to be aware of the slight variability in how these definitions are used to avoid misunderstandings. However, the majority of this review focuses on the most obvious and unambiguous definition of isogamy: syngamy and gametes exist, and only one gamete size is produced.
Regardless of the nature of isogamous reproduction (in a broad sense), it seems always to involve mating types—that is, genetically determined mechanisms that regulate compatibility between fusing gametes. The number of mating types varies; most commonly it is two, in which case the notation is typically + and −, but there can be more [51]: the social amoeba Dictyostelium discoideum has three [26,52], the ciliate Tetrahymena thermophila has seven [49,53–55] and fungi can have up to thousands of mating types [27] which, therefore, approach a continuum of compatibility types, and can make the vast majority of possible encounters sexually fertile [51]. Intriguingly, mating types and sexes can exist simultaneously and separately in one individual. For example, the haploid gametophyte phases of some fungi (heterothallic filamentous ascomycetes) are hermaphrodite, i.e. they produce both male and female gametes, but male and female gametes are only able to fuse if they originate from different mating types [56,57].

Superficially, mating types may seem similar to self-incompatibility in plants [58], and prevention of selfing is indeed one hypothesis for the origin of mating types. However, it cannot be a universal explanation, because some mating-type systems do not prevent selfing; if mating types are determined at the haploid level, meiotic products of diploid forms will inevitably have complementary mating types [56] and hence can fuse. Furthermore, some fungi are homothallic, meaning that they are, broadly speaking, self-compatible (see box 1). Saccharomyces yeast, for example, is capable of changing the allele at the mating-type locus, which means that a haploid form can mate with its own daughter cell, leading to a completely inbred offspring homozogous everywhere except the mating-type locus [28]. There are currently multiple hypotheses for the evolution of mating types, including a novel one in this issue [59]. These are beyond the scope of this review (but, e.g. [27,56,60–62] and references therein are useful starting points on the topic).

3. Evolutionary forces impacting isogamy
(a) How stable is isogamy, and when is it expected to become unstable?

Reasons for the male–female dichotomy are reviewed in Lessells et al., Togashi & Cox and Lehtonen & Parker [2,3,63]. To complement the emphasis of past reviews that consider explaining the shift towards anisogamy their main task, here we concentrate on the theoretically expected conditions under which isogamy should remain stable. Because these theories concern the evolutionary transition from gametes of one size to gametes of two sizes, this discussion is restricted to isogamy in the strict (gametic) sense, as discussed above.

The basis for the modern theory derives from a gamete-size model by Parker et al. [64]. Each parent has a fixed resource, $R$, to divide into gametes, implying that if $n$ gametes are produced then each gamete has the mass $m = R/n$. The size of the zygote resulting from fusion of gametes of size $m_i$ and $m_j$ is $S_{ij} = m_i + m_j$. Parker et al. simply assumed that all gametes from many parents with a sufficiently wide variation in $n$ fused at random in seawater. Briefly, if the ‘fitness’ $f$ of a zygote is an increasing function of its size, $f(S)$ (generally assumed to become asymptotic with $m$), there are two possible ESSs (evolutionarily stable strategies [1]): isogamy and anisogamy. Many subsequent analyses (beginning with [65]) start with + and − mating types, but this does not alter the two possible ESSs. This model yields conditions for the stability (or instability) of the isogamous state, depending on just two very obvious assumptions about the effect of gamete size on fitness (increased size decreases gamete number, but increases survival as a zygote). Maynard Smith [66] deduced that the one-sex, isogamous state would be an ESS if

$$m^* = \frac{f(2m^*)}{f(2m^n)} \quad (3.1)$$

Matsuda & Abrams [67] derived the same solution starting from pre-existing mating types. However, recall the definitions of an ESS and a convergence stable (or continuously stable) strategy: if selection drives the evolution of the trait towards a specific value, this value is convergence stable [68–70]. Once an allele coding for this value has reached fixation, if no other mutant allele can increase in number, it is an ESS [68–70]. Matsuda & Abrams [67] showed that equation (3.1) is an ESS, but not convergence stable (see also [1,66]).

So what, if anything, causes isogamy to be stable? This problem was first explicitly studied in a key paper by Matsuda & Abrams [67] (see also [71]), who investigated the evolutionary dynamics of isogamy in a population with or without mating types. They stressed the apparent rarity of isogamy and analysed a number of cases where the instability of the isogamous equilibrium (3.1) resulted in a shift to anisogamy. They proposed two central reasons why isogamy might remain stable. The first is that gamete-size genes fail to become linked to mating-type genes. In the ancestral isogamous population, size-determining loci are unlikely to have been linked to mating-type loci; this could prevent invasion if mutations with large effects on size are possible and viable. While this remains a possibility, it may be difficult to remain in a situation with no linkage: if at least some size-determining loci are linked to mating type loci, and gamete sizes of different mating types have become unequal, then selection acts to further reduce recombination between the gamete-size and the mating-type loci [65]. In anisogamous species, there is clearly tight linkage or sex-limitation of expression of gamete size-determining loci, a feature which will tend to block any return to isogamy from anisogamy [67].

The second potential reason for stable isogamy proposed by Matsuda & Abrams [67,71] is strong, direct stabilizing selection on gamete size (as distinct from zygote size), due to the influence of a gamete’s size on its ‘fitness’ (survival, or success at syngamy). Although Parker et al.’s [64] model used an arbitrary minimum gamete size, they argued that survivorship before fusion would constrain both the size of microgametes in anisogamous organisms, and the size of isogametes. Matsuda and Abrams demonstrated that including a function relating gamete size to its survival and fusion success could, under some parameter conditions, permit isogamy to be locally stable, with gametes evolving to be greater than an arbitrary minimum possible size. Under these parameter conditions, isogamy could exist as an alternative equilibrium to anisogamy, though other parameter conditions permitted only anisogamy (see also [72], which linked gamete success to motility).

Bulmer & Parker [73] used a game theoretical approach in models with separate zygote ‘fitness’ $f(S)$ and gamete ‘fitness'
(b) Isogamy is linked to unicellularity

There is a clear taxonomic link between isogamy and unicellularity. Although isogamy does occur in unicells (e.g. in some species of the unicellular green algae genera *Chlamydomonas, Carteria* and *Chlorogonium* [6]), it is less common than isogamy. In multicells the reverse applies: while isogamy does occur, most multicellular taxa—including all metazoans and angiosperms—show isogamy. It was Knowlton [38] who first demonstrated a link between increasing degrees of isogamy and number of cells in a colony in volvocine algae. Since then several other studies [6,39,80–82] have shown similar relations between isogamy and multicellularity in the Volvocales, though Madsen & Waller [82] found that oogamous species had smaller masses at maturity than might have been otherwise expected, an effect they linked to the ecological pressures of living in pools or other stressful conditions rather than lakes. Parker et al. [64] argued that increased investment in the zygote is likely to be much more favourable for multicells, because of their prolonged period of zygote growth and development in the build up to a much larger and more complex organism. They contrasted this with the case of many unicells with ‘gametes’ and ‘adults’ of similar orders of size, where the stable condition may be to produce the smallest gamete size that yields the maximum prospects of fusion. Good support that increased zygote size is associated with increased body size across taxa is given by Bell [81]. On similar lines, Matsuda & Abrams [67] proposed that isogamy should be more common in species whose gametic phase represents a larger fraction of the entire life history; smaller organisms often have shorter adult phases. This could also account for the association of isogamy with small adult body size (see also [6,82]). Cox & Sethian [72] predicted that gametes of isogamous taxa would generally have a longer period of fertilization competency than gametes of anisogamous taxa.

As we have seen, isogamy becomes unstable when the zygotic survival function, \( f(S) \), becomes separated from the gamete survival function, \( g(m) \). Bulmer & Parker [73] argued that for many unicells, \( g(m) \) and \( f(S) \) are likely to be similar in shape and location, keeping isogamy stable. Using the unicellular algae *C. reinhardtii* and *Carteria palmate* as examples, haploid parental cells differentiate into gametes, which fuse and form a zygote, which undergoes meiosis and again produces haploid cells (e.g. [34–36]). Because the gametes are derived directly from the mature parental forms, the two functions are unlikely to be very different.

In general, it is intuitive that when both adults and gametes are small, as in many unicells that produce gametes by only one or two divisions of the adult, then these functions are similar, as both gametes and adults do rather similar things. Complex multicellularity, on the other hand, may affect \( g(m) \) relatively little, but probably pushes \( f(S) \) to the right as the need to provision the zygote increases, causing isogamy to become unstable and generating the transition to anisogamy.

The likely effect of multicellularity on \( g(m) \) and \( f(S) \) can be envisaged mathematically using the Vance formulations of fitness in equation (3.2), relating to how provisioning (\( m \) or \( S \), which are in the same units) contributes to \( g \) or \( f \). Suppose

\[
g(m) = \exp\left(\frac{\alpha}{m}\right) \\
and \quad f(S) = \exp\left(\frac{-\beta S}{S}\right).
\]

which are simplifications of forms derived by Vance [74,75] and Levitan [76] for survival of gametes and zygotes in marine invertebrates. They generate sigmoidal fitness curves as mass \( (m \text{ or } S) \) increases. Isogamy is stable provided that \( \beta < 4\alpha \). When \( \beta \) increases beyond this value, there is a switch to anisogamy. We can also reasonably assume that \( \alpha \) cannot be much smaller than \( \beta \), given that at the limit we expect the gamete and zygote to be of roughly similar size. So the stability of isogamy depends on the two survival functions, \( f(S) \) and \( g(m) \), remaining rather similar, with the rise from minimum to maximum survivorship occurring at roughly comparable levels of provisioning. This can be related to biological complexity—the two curves are likely to separate on the provisioning axis as multicellularity develops (see section below on the link with unicellularity). All the above models are based on the assumption that fertilizations occur in large pool of gametes shed by a large population of parents. Lehtonen & Kokko [77] have shown that isogamy can also be stable if one deviates from this picture and limits the number of adults that contribute to the total gamete pool (for other models involving low gamete competition, see [78,79]). Their model varies the size of the local group of ancestral broadcast spawners, explicitly ensuring that proto-male and proto-female fitnesses are equal (the ‘Fisher condition’), which is especially important when considering small groups of adults. They used the same form for \( f(S) \) as in (3.2), but used a related form for \( g(m) \) that more realistically modelled survivorship before fusion, by allowing this to increase with gamete size and by also tracking the time it takes until fusion occurs (this time is longer if fusion opportunities are limited). For high gamete competition, their results matched those of Bulmer & Parker [73], but the opposite assumptions (few adults in the local mating group) led to a new reason for isogamy to be stable. If the group consists of just one individual of each mating type, isogamy can be favoured since both individuals do best when they act jointly to maximize the number of viable zygotes; cooperation is maintained even though each parent is, of course, acting in its own best interests. This solution, however, requires that gametes find each other without too much difficulty. If they do not (e.g. the parents are far away from each other), isogamy is no longer a ‘waste’ of a large numbers of gametes which die once all the macrogametes have fused; it is now in both parent’s interests that one specializes in producing numerous ‘searchers’.

When local interactions involve more than one individual of each mating type, competition between individuals of the same mating type can destroy the above ‘cooperative’ nature of the game, and microgamete size evolution is mainly driven by competitive interactions. However, isogamy does not become immediately impossible as soon as more than two individuals interact, though it does become less likely with increasing gamete competition. In Lehtonen & Kokko’s model, conditions for isogamy to be stable become more permissive than found by Bulmer and Parker; \( \beta > 4\alpha \) is still a necessary, but not sufficient condition to guarantee the switch to anisogamy.
that the size at which the gamete and the zygote reach similar ‘fitnesses’ (i.e. where $g = f$) is $G$ for the gamete and $A$ for the zygote (while noting that the interpretation of ‘fitness’ for a gamete has to be done within a limited scope: we talk about success up to the next stage in life in a sequential model). To a rough approximation, $G$ and $A$ can be related to the typical mass of the gamete and zygote at the end of any growth stage: $f(S)$ is the contribution made by the provisioning, $S$, to the future fitness of the zygote at (say) sexual maturity. Then $e^{-u/S} = e^{-p/A}$, and hence $A/G \approx \beta/\alpha$. For many complex multicellular organisms, $A \gg G$, which is equivalent to asserting that $\beta \gg \alpha$. Note that this explanation is heuristic only; to expect the two curves to be the same shape is a simplification, since the two stages in complex multicells typically have radically different life styles (e.g. gametes may not feed or grow, while zygotes usually do) and expected lifespans. Unlike their unicellular counterparts, zygotes of multicellular organisms give rise to embryos that often grow vastly in mass (and cell number) before sexual maturity, so the contribution due to zygote provisioning $S$ to future fitness becomes small as $S$ becomes a trivial component of the total mass. Even so, being a large zygote may be beneficial for a multitude of reasons, from energy advantages of being well provisioned to having outgrown some predator attacks from the outset (e.g. protists are important predators of planktonic life) to predation itself being easier if the predator does not have to grow from a small egg—this can even impact clearly differentiated developmental stages of fish larvae that cannot feed efficiently if too small, because of hydrodynamic properties of water itself [83].

The optimal gamete size relates to surviving to successfully fertilize another gamete, and does not depend in an obvious way on the size of the mature individual. But the optimal size for a zygote relates to surviving until maturity/age of first reproduction, which almost inevitably does depend on the size of the mature individual. So it is likely to be organismal mass and complexity, rather than multicellularity per se, that generates the separation between the $g$ and $f$ functions.

Although a loose link between unicellularity and isogamy exists, it is by no means perfect. What about the exceptions? A full investigation will require empirical constructions of the $g(m)$ and $f(S)$ relations for several specific cases, including those that do not fit the link. A full theory will also keep mindful of the biological diversity: in *Chlamydomonas*, nitrogen limitation induces gametogenesis, but gametes can grow back (dedifferentiate) to vegetative cells if the availability of nitrogen—and thus the conditions for vegetative growth—improves ([84,85], J. Tomkins 2016, personal communication). This is a case of extremely flexible facultativeness of sex, as even once a gamete has already been produced, there is no firm commitment to sex. This is possible because in a *Chlamydomonas* life cycle, a sexually produced diploid zygote produces ‘adult’ vegetative cells via meiosis, and gamete production thereafter does not require another meiosis. Because vegetative cells and gametes do not differ in ploidy, there is nothing preventing a gamete from growing back to its former state should this be favoured by current conditions. Conceivably, anisogamy would in such a system have an additional cost that is not included in current models: very small gametes might have to forego the (presumably selectively advantageous) flexibility and commit to finding a partner to fuse with, if the strategy of growing back to adult unicell size is no longer feasible from very small sizes.

(c) The difficulty of returning from anisogamy back to isogamy

We have seen reasons for the instability and taxonomic distribution of isogamy, but there is another important and related theoretical prediction: once anisogamy evolves, it is expected to be very stable [86]. The reasons for this have been discussed extensively elsewhere [63,77,86]; briefly, the main point is that when gamete dimorphism has evolved to a significant level, an extremely low level of sperm competition (or more generally, gamete competition) is sufficient to maintain it. This is intuitively understandable. As soon as gamete sizes have diverged and one gamete type outnumbers the other, many gametes of the more numerous type (e.g. sperm) are destined to remain unfertilized. Because it is not known in advance which male gametes will be successful, a relatively small increase in the reserves of the few successful sperm would require wasting the same amount of extra provisioning on a large number of unsuccessful sperm. On the other hand, even a relatively large change in the size of a tiny sperm size can still be very small in relation to the size of the egg, therefore, contributing little to the survival of the zygote. The result is that increasing zygote provisioning by a significant amount requires a large decrease in microgamete numbers, with a corresponding decrease in sperm competition ability. A return to isogamy could theoretically happen if either the zygote provisioning requirements decreased (e.g. due to an evolutionary reversal to unicellularity), or if both sperm competition and sperm limitation were entirely absent [63,77]. Another factor that makes a reversal to isogamy difficult is that anisogamous species have achieved tight linkage or sex-limitation of expression of size-determining loci, which can block the evolutionary path from anisogamy back to isogamy [67]. Despite these stabilizing factors, there are known cases where ‘giant sperm’ have evolved secondarily [87–89], approaching isogamy in terms of size; such special cases are, however, quite different from an actual evolutionary reversal to morphological similarity, because a clear sperm–egg dichotomy in morphological traits is maintained.

4. The consequences of the transition from isogamy to anisogamy

As we have seen, there are theoretical reasons to expect isogamy to be evolutionarily unstable in complex, multicellular organisms [67,73,77]. This has some fundamentally important consequences beyond the obvious one of having two different gamete sizes.

The first is really a definitional one: the transition to anisogamy implies the appearance of male and female sexes (or male and female functions in simultaneous hermaphrodites). If we take the definition literally, males and females essentially ‘appear’ as soon as the slightest consistent gamete competition (or more generally, gamete dimorphism) evolves. Furthermore, as we have seen, it is difficult to go back to isogamy once this divergence happens.

Second, as soon as the disparity between male and female gametes increases further, some additional consequences follow. It has traditionally been thought [90,91] that the
dimorphism in gamete sizes also tends to lead to dimorphism in male and female behavioural traits (sex roles). For example, males tend to be more competitive for fertilization opportunities than females (e.g. [91,92]). Theory suggests that this can be driven by the numerical imbalance that results from anisogamy: the producers of the more abundant gamete type have more scope to increase the fertilization probability of their gametes, which can select for such ‘competitive’ traits at both the pre- and post-ejaculatory stages [20]. While the authors of this article are in agreement with the notion that typical male and female sex roles derive from the initial sexual asymmetry of anisogamy as originally proposed by Darwin [90] and Bateman [91], we must note that this view has recently been the subject of debate (e.g. [93–95] against, [19,20,96,97] for). More generally, we would argue that most phenomena usually considered to be sexual selection and sexual conflict ultimately arise from anisogamy. However, as discussed elsewhere in this theme issue [98], sexual selection does not strictly require anisogamy. Nor are isogamous organisms necessarily free of all sexual conflict: a theoretical prediction suggests that total investment into a zygote (i.e. sum of gamete sizes) under isogamy can sometimes be less than it would be under anisogamy (or asexual reproduction) in similar conditions [67,71,77], resulting in decreased zygote survival (see, e.g. fig. 1 in [67] and fig. 5 in [77]). This has been interpreted as a theoretical example of conflict between the ‘proto-sexes’, where each parent would benefit from the other investing more in a similar manner to conflict over parental care [99].

Finally, a major consequence of anisogamy is the ‘twofold cost of sex’ [66], which arises when gamete dimorphism evolves to the point that male gamete provisioning to the zygote is negligible compared to females. This implies that if an asexual (female) mutant that does not require fertilization and does not produce males appeared, she would be able to produce female offspring at twice the rate as her sexual counterpart, and the same would continue in the offspring generation and so on. This is the twofold cost of sex, or perhaps more accurately, twofold cost of producing males in a gonochorist. It is one of the prime reasons why the maintenance of sexual reproduction is considered such a mystery (e.g. [66,99,100]; see also the latter two for cost of sex through ‘genome dilution’ in hermaphrodites). The crucial point here is that the twofold cost only appears with the evolution of anisogamy, and does not affect isogamous organisms, for the following reason: although sexual reproduction takes place in isogamous organisms, there is no ‘parasitic’ gamete producer in hermaphrodites. The crucial point here is that the two-female mutant that does not provide any developmental resources to the zygote (assuming no paternal care). Both parents of every offspring provide resources, and every offspring is in turn capable of providing parental resources to their own offspring. While males are a ‘cost’ in isogamous systems, in isogamous systems there is no such costly type that provides no developmental resources to the zygote.

As we will see next, this has some interesting consequences for the evolution and maintenance of sexual reproduction: can we learn something about sexual reproduction by studying both isogamous and anisogamous species?
of sexual and asexual reproduction in isogamous and anisogamous organisms. The aim is to bring an alternative perspective into the debates surrounding the origin and maintenance of sex itself. Instead of looking for a higher than twofold benefit of sex, we ask what would happen if reaching high short-term benefits was rarely the case. Although both long- and short-term benefits exist (as is very plausible based on numerous studies [101]), the latter is assumed to be generally lower than twofold in magnitude. We then combine this assumption with the evolutionary history and most probable ancestral states of sexual reproduction, and examine the patterns that are expected to be generated over evolutionary timescales.

Three points are particularly important here. First, the twofold cost of males applies only to anisogamous species (see above; [66,99,100]). Second, anisogamous eukaryotes have likely evolved from isogamous ones [2,3,73,77], which were hence free of the twofold cost. Third, the ancestral reproductive strategy in eukaryotes was facultative sexual reproduction [9,10].

Consider then an ancestral organism, with facultative, isogamous sex. Facultative sex is known to be potentially as beneficial as obligate sex [102]. Additionally, facultativeness and the absence of anisogamy both help to keep costs small. Thus, sex is not hard to explain at this stage [103,104]. Because asexual life cycles occur regularly in species with facultative sexual reproduction, adaptations that keep the asexual life cycle efficient are maintained by natural selection. As we have assumed that sex is advantageous to some (lower than twofold, for reasons that will become apparent later) degree, and not very costly in the absence of males, a subset of isogamous species may turn facultative, and become reliant on sex to such an extent that asexual reproduction is unlikely to invade success-fully [106]. This makes the rising cost of males irrelevant: it becomes very costly. Group C—species with long established obligate sexual reproduction with strong constraints that make the re-invasion of asexual reproduction extremely unlikely.

Anisogamy can start to evolve (i.e. gamete sizes can start to diverge) in any of these three groups. Regardless of the ultimate reason why anisogamy evolves, it shifts the system to the derived high-cost state: an increasing degree of gamete dimorphism predicts an increase in the cost of males over time.

With our assumption of lower than twofold short-term benefits of sex, the effects of an increasing cost of male production differ between the groups. Group A is a priori capable of efficient asexual reproduction, making it easy to respond to an increasing cost of sex by engaging in the sexual cycle less often. Species in group B can relatively easily regain the lost capacity for asexual reproduction under increasing costs of sex. If a reversal to asexual reproduction appears, it has a high probability of invading in these two cases as the cost of males can be near twofold. As we have seen, anisogamy itself is unlikely to reverse. Groups A and B are, therefore, selected to re-evolve obligate asexuality: males and the capacity for sex are purged as sex becomes very costly. Group C is the only one in which anisogamy can invade such that sexuality remains unaffected. This group has adapted to and become reliant on sex to such an extent that asexual reproduction is unlikely to invade successfully [106]. This makes the rising cost of males irrelevant: it can increase to twofold (or beyond [99]) and sex can still be maintained.

While it has been pointed out before that the evolution of sex may be influenced by constraints [102,107,110–113], this connection between facultative sex, the evolution of anisogamy and the associated changes in the cost–benefit balance of sex appears underappreciated. Once this connection is taken into account, it is no longer mysterious why sex persists even though obligate, constrained sexual reproduction may lead to lower individual fitness in anisogamous species than would a successfully performed asexual life cycle. It is the only form of sexual reproduction where anisogamy makes it highly costly, we need to understand that anisogamy could, at least in principle, invade very different isogamous organisms. We assign isogamous organisms into three main groups as follows (figure 1):

Group A—ancestral species with facultative sexual reproduction, which can be assumed to be capable of efficient asexual reproduction;
Group B—recently derived species with obligate sexual reproduction, with no constraints on the re-invasion of asexual reproduction, or constraints that are relatively easy to overcome;
Group C—species with long established obligate sexual reproduction with strong constraints that make the re-invasion of asexual reproduction extremely unlikely.

Anisogamy makes it highly costly, we need to understand that this transition can be a rare event without affecting the argument. Any cost is now paid every generation but there still is no male cost. Evolutionary reversals to asexuality are possible in this group if the selective environment changes.

However, lineages that have been obligately sexual for a long time tend to accumulate constraints that make invasion by asexual reproduction unlikely. Derived asexuality is often associated with fitness problems reflecting a legacy of past sexual reproduction [106]. Thus, a multitude of processes decrease the likelihood of the invasion by asexuality after the lineage has relied on obligate sexual reproduction for a long time [107]. Many examples of such constraints are known in multicellular organisms [107]. For example, gamete activation may be required, where sperm (either from the same or a related species) is needed to trigger egg development even if there is no genetic contribution from sperm (‘sexual parasitism’ and gynogenesis in particular; reviewed in [108]); reduced offspring fitness via inbreeding depression can prevent the spread or maintenance of some forms of asexual reproduction which increase offspring homozygosity [107]; and in some species, only sexually produced eggs are resistant to harsh conditions, which may act as a lineage-specific mechanism to maintain facultative sex [107,109].

While such constraints are best known in multicellular anisogamous animals, many of the mechanisms could equally well apply to isogamous species. In isogamous cases, such constraints can have ample time to evolve because sex is not very costly (and thus more easily maintained) before anisogamy has evolved.

To explain why constraints, then, can play a role in the very large number of cases of sexual reproduction where anisogamy makes it highly costly, we need to understand that anisogamy could, at least in principle, invade very different isogamous organisms. We assign isogamous organisms into three main groups as follows (figure 1):
in any group, but the emerging males disappear as asexuality reinvades and sex is lost—unless sexual reproduction has been in use for long enough to have become protected by significant constraints. The temporal order of evolutionary transitions is thus important in this thought experiment.

What does this predict for the phylogenetic distribution of sex? The derived state (right side of figure 1) suggests a specific evolutionary pattern. Groups A and B are rare once they have reverted to asexuality, based on inferior evolutionary success over long timescales [66,111,114—119]. Group C is the only one that avoids these adverse long-term effects after gamete sizes diverge: in spite of short-term costs of sex, the long-term benefits have ample time to operate because constraints prevent these lineages from reverting to asexuality. After anisogamy has evolved, sexual selection can reinforce the constraints [112] making a return to asexuality even more difficult. At the same time, sexual selection and conflict increases the potential for ecological diversification [120]. Thus, a relatively small number of group C species passing through the ‘anisogamy gateway’ can suffice to form the ancestors of the diversity of anisogamous life we see today: the pathway from the top left to bottom right in figure 1 can be rare.

Despite anisogamous species in group C not reverting easily to asexuality, the large number and high diversity that they reach over time makes it possible for asexual reproduction to occasionally re-invade some lineages (bottom right corner of figure 1). This makes the low transition rate predicted by our account compatible with the observed ‘twiggy’, sporadic pattern of asexuality around the tree of life [121]. As long as there are ecological differences between sexual and asexual strategies in a single species, this is also compatible with the observed cases of facultative sex in anisogamous species. Such differences can maintain both reproductive modes in a single species [122]. Examples include limited dispersal ability of vegetatively produced progeny compared with outcrossed seeds in many plants, or the dormant sexual zygotes of aphids [122,123] and rotifers [109].

Faculative sex can be derived from obligately sexual anisogamous ancestors (e.g. aphids [123]). Alternatively, if ecological differences arise before the evolution of anisogamy, facultative sex may remain stable despite the evolution of anisogamy. This could explain the existence of groups of facultative sexuals with both isogamous and anisogamous species, such as volvocine algae [124].

Although this is a somewhat speculative thought experiment, our conceptual model is quite general in its assumptions. It does not rely on a specific benefit of sex; we simply assume that short- and long-term benefits exist. The only requirement is that the former must be lower than twofold in magnitude. It also does not rely on any specific explanation for the origin of anisogamy, nor on the precise type of constraints that keep group C sexual. However, one possible caveat is that the mechanics of meiosis can pose a significant time cost for unicellular organisms [99,125,126], and the exact magnitude of these costs is of course important in determining how easily obligate sex can actually arise in an isogamous ancestor.

This conceptual model makes some clear predictions. The majority of strongly anisogamous organisms are predicted to be obligate sexuals, and to have very low transition rates back to asexual reproduction. This could explain the maintenance of sexual reproduction in anisogamous species, and is in line with previous work and patterns seen in nature [107,121].

Another testable prediction of the conceptual model is that facultative sex should be more common in isogamous organisms or those with a low anisogamy ratio than in strongly anisogamous species (compare left and right sides of figure 1). For this, there seems to be tentative support in the literature, either directly [127], or via the association between multicellularity and anisogamy [10]. However, for a large proportion of microbial eukaryotes we do not yet know if they ever reproduce sexually [9,128], and many of these are not strictly gametic systems (see above). Theory on the evolution of anisogamy predicts isogamy to be common in small eukaryotes of low complexity [73,77], and facultative sex is known to be prevalent in microbial protists [10,102], which gives further (albeit indirect) support for this prediction. Our results stress the importance of understanding the frequency of sexual and asexual life cycles in microbial eukaryotes [9], and of quantifying not only benefits but also the varying costs of sex [99], to piece together the big picture of the evolution and maintenance of sexual reproduction. Many enigmatic features of sex may become explicable once one allows for a changing cost of sex across evolutionary lineages, with a particular emphasis on both sides of the transition to anisogamy.

6. Conclusion

Of all the eukaryotic diversity, we quite understandably know much more about multicellular organisms than unicellular ones (e.g. [129]). For example, although the ancestor of all extant eukaryotes was likely facultatively sexual [10], for a large fraction of unicellular organisms we still do not know if they reproduce sexually or not (although facultative sexuality in microbial eukaryotes seems more common than is immediately apparent [9]). The transition to multicellularity is clearly a major event in eukaryote evolution, partly because it seems to be linked to the evolution of anisogamy [37], to the costs of sex [99] and, seemingly paradoxically, to obligate sex [10].

Although this article is focused on isogamy, our main message is more general: that even if our focus is on understanding multicellular organisms like ourselves, some of the most fruitful avenues to this may be found by also studying organisms that are very different from us. Our thought experiment in the last section is intended to exemplify this: although it is speculative, it gives an alternative viewpoint on the evolutionary history of sex, and shows that there is a plausible pathway by which sex could have become an almost irreversible feature of multicellular life, even if it does not come with strong short-term benefits. Seeing this pathway requires looking at both isogamous and anisogamous organisms in tandem.

Competing interests. We declare we have no competing interests.

Funding. J.L. was funded by a University of New South Wales Vice-Chancellor’s Postdoctoral Research Fellowship and a University of New South Wales Early Career Research Grant.

Acknowledgements. We thank Madeleine Beekman and three anonymous reviewers for constructive comments that improved the manuscript.


