Dear Dustin,

Thank you and the other two reviewers for your very constructive comments and criticisms of our manuscript. Learning what was not understood has been illuminating and has led to a complete re-write and re-organisation of the manuscript. Additionally, the revised version now includes new explanatory figures and a mathematical model. This has all taken much longer than intended, so thank you for your patience.

Arguably the biggest source of confusion with the previous version was (previous) Figure 2, where we had attempted to draw attention to the key difference between the two (mixed and non-mixed) treatment regimes. In light of the reviewers' comments it is clear that the figure failed to convey our message and gave the impression that our experiment involved a manipulation akin to a trait group model. This is not what we did.

We have redrawn Figure 2 – it is now Figure 1 – to make, we hope, abundantly clear that the difference between treatment regimes is simply whether or not propagules experience mixing during the dispersal phase. With this clear, we think that much that was previously confusing has been rectified.

To take a step back: this paper concerns ecological conditions fuelling evolution of a simple life cycle. In our previous (2014) paper, nascent multicellular lineages, comprising soma- and germline-like phases, engaged in a birth-death process without ever mixing. Our paper compares this treatment with one in which we maintain independent nascent multicellular lineages, but additionally allow competition (by mixing) among the germline-like phase. This seemingly insignificant alteration has marked effects on the evolutionary fate of the evolving collectives.

In this regard the paper is far more than a follow-up of our previous 2014 study. It shows clearly that ecology matters and it leads to some exciting conclusions regarding factors that affect how selection sees nascent lifecycles. In revising the paper we have sought to bring to the fore these new insights.

An issue raised in two of the reports was absence of a discussion of kin selection or use of inclusive fitness to interpret our findings. We think this is largely a consequence of our failure to make clear what we actually did – and which we hope is rectified by the changes mentioned above. But additionally, we have changed the names of the two life cycle phases from "Phase I" and "Phase II" to "Maturation Phase" and "Dispersal Phase". The Maturation Phase is identical for both treatments and starts with a single-celled bottleneck every life cycle. This means that relatedness within each group is the same: regardless of treatment, $r = 1$.

Both treatments underwent evolution in the same structured ecology during Maturation Phase. It is during this phase that we took our measurements of group fitness, cell fitness and other parameters, in part to control for any differences in confounding factors such as variation that may occur during the Dispersal Phase.

The only difference between the two treatments, as above, occurs during the Dispersal Phase, where propagules are mixed from all groups in the Mixed Treatment, and from each group separately in the Non-Mixed treatment. We have also re-named the Mixed Treatment "Mixed Propagule Ecology", so that readers are reminded of this subtle distinction throughout the manuscript. The relatedness between cells in the Dispersal Phase is not relevant to the inclusive

fitness of the multicellular groups, in the same way that the relatedness of competing sperm from different fathers has no bearing on the inclusive fitness of cells in their offspring.

It is of course quite probable that there are differences in the level of diversity present during the Dispersal Phase. Given our previous results indicating an increase in mutation rate in the Non-Mixed treatment, it is possible that there is even more diversity (less relatedness) within the Non-Mixed groups than groups in the Mixed treatment. We do not discuss this here, because as you have rightly pointed out, we need to be careful concerning our aims and experimental design. Our paper is already conceptually dense, and given that we did not set out to examine diversity of dispersal cells, we do not wish to deviate from the main points of the study.

We show that our findings are due to a trade-off, originally present in the ancestor, between the traits that underpin group fitness (Transition Rate), and cell fitness (# of WS cells). We have now included a mathematical model made by Yuriy Pichugin (now a co-author), which supports these findings, and shows that if the trade-off were to be removed, we would have very different results. Importantly, the model is consistent with our findings and does not rely on arguments concerning relatedness.

Further apparent from the referees' comments is that the summary figure (previously Figure 7) was difficult to understand. We have redesigned this figure (now Figure 9), including only our main results and conclusions, emphasising the combination of the underlying trade-off and ecology as the responsible for our findings. This ecological structure lays the foundation for subsequent selection of a developmental program, and with this the establishment of kin-groups and thus the possibility for a kin-selection-like process to occur.

Below we respond in detail to the reviewers' comments and hope that both you and the reviewers will find our re-worked manuscript much improved.

Thank you again for the encouragement to strengthen our paper, and we look forward to hearing your decision.

Yours sincerely

Caroline Rose and Paul Rainey (on behalf of the authors)

Reviewer 1

This is an interesting set of experiments that I think can have an impact if it were modified in a few ways. One of the most important is the presentation of the work. The writing is difficult to follow because it lacks some structure and is uses poetic language where more direct language would be more clear. More importantly, it is not always clear what has been done and what it means. A general re-write that makes the descriptions more concise and direct would dramatically improve the audience breadth that could be impacted by this article.

Point taken. We have taken this on board and rewritten the manuscript, re-made two of the figures and included a mathematical model to draw attention to the fact that our experimental findings are consistent with expectations from a simple model.

My understanding of this work is that the experimental setup forced near-complete group selection in one scenario and ONLY individual selection (selection between groups cannot operate if only the most successful individual within any group seeds the next generation) in the other.

Selection in the Mixed treatment was implemented between groups during the Maturation Phase, and between cells during the Dispersal Phase, leading to an interesting conflict between the two life cycle phases, due to an inherent trade-off. The differences between the two treatments, as well as the two life cycle phases within treatments, has been made much more explicit.

What was less clear was the major conclusion. Some parts of the writing (see the final paragraph) make it seem like the point is to demonstrate that group selection can have an impact on evolution, but that is well established at this point. I think the major conclusion is that group selected lines had better group fitness (although I am not entire clear how group fitness was measured). What is missing is a why; why are those with better group traits less likely to go extinct? I am pretty sure that it is bc there is no transition in the time frame, but it is never clearly described. I think this is mostly a presentation issue.

The trade-off between traits that are selected at the cell- and group- levels is now fully emphasised in the revised manuscript. We also make transparent the fact that both traits cannot be maximised, and show that in the Mixed Propagule Ecology, cell-cell competition in the Propagule Phase ‘triumphed’ over between-group selection during the Maturation Phase. This led naturally to our conclusion that evolutionary transitions in individuality require groups to be discrete throughout their entire life cycle (at least in the early stages). This is likely to differ from many interpretations of standard ‘group-level selection’.

Review 2

1- It is somehow difficult to follow through the text from Introduction to Results without relevant methods explained or mentioned.

We hope that the revised and substantially re-written manuscript solves this matter.

Line 178-180 (also line 552-553), it is not clear to me how group fitness was measured, especially given that how it was measured is important to interpret the results. Does “ability to

leave group offspring" mean the proportion of cells that became "WS " cells during 3-day period in Phase II after being plated out?

Our group fitness assay is now explained fully in the Results and Methods sections: "ability to leave group offspring" is not the proportion of cells that became WS after the Dispersal Phase (II), but is the proportion of groups (out of 8 groups in a metapopulation) that are "won" by the strain being assayed, relative to the reference strain, after one entire life cycle. The single-cell bottleneck in both treatments ensures that there are no chimeric offspring groups, so each offspring group can be assigned a "winner".

Line 174-176, could the authors explain more on how competition assays were conducted? For example, are marked ancestral lines WS or SM cells? Were evolved lines competing against their own ancestral lines or a common ancestor?

Thank you for the suggestion. We now explain in the manuscript that the marked ancestral (reference) strain is a SM, and that we also generated a range of "ancestral" WS strains, just one mutational step (or half a life cycle) from the true ancestral SM strain. This range of ancestral WS strains enabled us to compare the full range of ancestral phenotypes and their relationships to each other (i.e. the shape of the trade-off that underlies the direction evolution will take), with those of the evolved lineages. All evolved lines, as well as the range of ancestral types, were competed against the marked reference strain.

2- Paragraph line 250-260 is a very insightful discussion on why the density of WS cells decreased in the Non-Mixed Ecology but increased in the Mixed Ecology. If I remember correctly from Hammerschmidt et al (2014), the transition between WS and SM cells in the Non-Mixed Ecology was mainly achieved through a mutS-dependent switch in wspR. This genetic composition may limit the possible WS cell types that may arise and therefore lower the competition among WS cells. I was wondering that whether such a switch evolved in the Mixed Ecology. If not, there might be more mutational space to explore and evolve different WS cells especially the ones with higher growth rate.

This is an interesting suggestion. It is true that the increased rate of transition in one of the Non-Mixed lineages is due to a *mutS*-dependent genetic switch. This is however, only one of 15 lineages in the Non-Mixed treatment, and many of the other lineages still transitioned between WS and SM by 'regular' mutation. We did not find such a switch in any of the Mixed Ecology lineages, and it is possible that their higher growth rate resulted in more WS-space to explore. However, we suspect that the reduced Transition Rate in this treatment is most likely due to a reduction in overall rate of mutation, and therefore possibly explore less mutational space.

Review 3

General comments

Firstly, there are a few experimental points that I think need clarifying. I should note here that I do not feel I have the expertise to comment in detail about the experimental methods, but perhaps these observations could lead to more clarity in the manuscript.

A key question for me is whether the non-sticky 'germ' cells have to arise de novo every generation? (For it to be considered a true germline, these cells would need to arise through a

developmental program, and not through random mutation and loss of the 'sticky' gene). Could the authors please clarify.

This interpretation is correct - the 'germ' cells do have to arise de novo every generation. We now make this explicit in the manuscript. Agreed – the development of other cell types must eventually happen through a developmental program, so here they are not a 'true' germ line, just as the bacterial groups in our experiments are not true multicellular organisms. We approach the evolutionary transition to multicellularity from the bottom-up. We ask how such a developmental program could come to be selected in the first place, from the starting point of cooperating groups of cells that do not have a developmental program. Here we take the negative-frequency-dependent oscillation between two cell types, by mutation, and explore the ecological conditions necessary for selection to act on this cycling itself (the longer timescale) rather than on just cell growth in the short-term. The idea is that fine-tuning of the cycling (by a developmental process) will emerge if ecology allows selection to act on the longer timescale of the life cycle itself. The introduction has been rewritten and these ideas are now explained much more clearly.

Another key point is whether the cells in these life cycles can survive on their own. A key feature of obligate multicellular organisms is that they are unable to survive and reproduce outside of the multicellular body – however it seems to me that these cells (of both types) can function outside of this life cycle. Could the authors please explain this more.

See above. Other groups studying the evolution of multicellularity take the "top-down" comparative genomics approach to studying simple multicellular organisms and discovering genes or traits essential to the transition. Here we take a "bottom-up" approach because we are interested in the selective and ecological causes of multicellularity, from the standpoint of cooperating groups of cells (single cells, from a selective point of view).

I would like to highlight that it seems unlikely that this specific experimental setup reflects 'real-life' scenarios that could have occurred at the beginning of the evolution of multicellularity.

We agree absolutely, and address this point in the Results and Discussion section. We suggest that the severe constraints on the ecological conditions required for a true evolutionary transition in individuality explains why cooperating groups of cells, while extremely common, do not, in general, evolve to become multicellular organisms.

I wonder if they could give some more examples of where these types of life cycles could be found in extant species?

Almost all animals reproduce via the "Non-Mixed" life cycle, although we do discuss the possibility of relaxing the condition of a discrete dispersal phase after the evolutionary transition to multicellularity was complete, eg. in the case of sperm competition in polyandrous animals.

Secondly, I would like to raise several general theoretical points in relation to possible misunderstandings of social evolution theory.

It has been shown that obligate multicellularity (= individuality) has only arisen when multicellular groups are formed through cells sticking together after division, thus guaranteeing that relatedness between them = 1 (Fisher *et al*. 2013). This is relevant, because (as far as I

understand it) the experiments in this manuscript essentially produce genetically identical (non-mixed ecology) and non-genetically identical (mixed ecology) treatments, where group formation is being experimentally manipulated.

We agree that obligate multicellularity must arise when multicellular groups are clonal. This is exactly why we have placed such importance on the single-celled bottleneck at the start of the Maturation Phase **in both treatments**. Group formation is not experimentally manipulated between the treatments – in both treatments, groups are formed and undergo Maturation identically. What is experimentally manipulated is the amount of competition between groups during the Dispersal Phase only. We accept that this misunderstanding is due to a failure on our part to be clear, and we hope that this has been remedied in the new manuscript.

Therefore, from what I can see, the experiments support the idea that clonal relatedness is important – which has been shown many times in a very strong collection of papers (most crucially the work of Ashleigh Griffin & Stuart West). This literature is notably absent from the bibliography.

See above. Yes, we agree that clonality is important and has already been shown many times, which is why we don't address that question here. The set of literature mentioned is not cited, not because we do not agree with it, but because we do not experimentally manipulate relatedness here – it is not relevant to our question. We were very careful when designing the experiment to ensure clonality in all treatments and only to manipulate the ecological structure.

The authors imply that group selection is "often dismissed as a rare occurrence" (line 508). This is a misunderstanding of social evolution theory, which I think it is extremely important to clarify. From West *et al*. 2015, Fig 4 legend: "Natural selection will lead to the evolutionarily stable strategy (ESS), which will be the strategy that maximizes inclusive fitness, irrespective of the consequences at the group level. We would expect natural selection to lead to maximization of group fitness, and thus think of the group as a fitness- maximizing individual, only in extreme cases where there is no within-group conflict". This is NOT saying that group selection doesn't happen in nature, it is in fact saying that it is always individual level inclusive fitness maximization, which appears as/is equal to group- level selection BUT ONLY WHEN the individuals within the group are without conflict (and this is most likely when relatedness = 1). In other words, inclusive fitness is always maximized, and sometimes (when r = 1) this coincides with maximizing group fitness as well, and this is when we see group adaptation.

We have removed this part of the Discussion as it was really not necessarily. Nonetheless, we disagree on this point, but suspect that it is possibly due to semantic confusion, in particular with regards to the definition of group fitness used. Group fitness is often interpreted as the average cell fitness. When interested in social evolution theory (the evolution of cooperation, for example) this definition of group fitness can be useful. However, this paper focuses on evolutionary transitions in individuality, and the ETI literature has a different definition of group fitness. We are interested in knowing when a group of cooperating cells becomes an evolutionary individual, that is, an entity that is capable of evolving by natural selection that is **not** a byproduct of selection acting at the cell level (or maximizing inclusive fitness). Group fitness in discussions of ETIs (and the one we use in this paper) is therefore defined as the ability of a group to leave group offspring.

Even when using this definition, increased group fitness can be interpreted as being due to selection acting at the group level (and therefore the group can be seen as an evolutionary 'individual') ONLY when group fitness is NOT caused by increased cell fitness. This is the

opposite of the scenario you describe above: "when individuals within the group are without conflict, inclusive fitness is always maximized, and this coincides with maximizing group fitness as well, and this is when we see group adaptation". In definitions of group fitness from the ETI literature, this is in fact not the case, because group fitness is simply a by-product of increased cell fitness. In other words, true group adaptations leading to an ETI are those that cannot be boiled down to the lower level (by definition).

For further interested, we discuss this in Hammerschmidt et al (2014), where we had a treatment exactly like the scenario you describe above. In that treatment, there was a single-celled bottleneck (i.e. r = 1) but there was no cycling between SM and WS cells, the cooperative WS cells were chosen during both phases of the life cycle. As expected, conflict was reduced (the transition rate to SM decreased), and group fitness increased. Cell fitness also increased, and we showed that the increased group fitness was a by-product of cell-level selection. Such a selection regime will never lead to a transition in individuality.

Detailed comments

Line 33-35 – sentence doesn't read well and needs editing

The abstract has now been completely re-written.

Line 55 -56 – this implies that the lower-level units didn't (or don't) 'participate' in evolution by natural selection. Both the lower-level units and the higher-level units are subject to natural selection, and I think that needs to be made clearer here.

We did not mean to imply this. This sentence is now gone.

Line 58 – 59 – I disagree with this statement. Some of the explanations for the evolution of multicellular life centre on the origins of group reproduction, but not all of them by a long way. Some of the explanations centre of suppression of cheating, some on the mode of group formation, others on the emergence of complexity and division of labour.

The introductory paragraph now includes other explanations.

Line 60 – 61 – I don't quite understand this statement, but I think it is saying that the somatic part of the life cycle is the part which experiences selection and the germline does not? Needs to be made clearer in order to be understood.

Again, that is not what we meant to imply. The importance of timescales is now explained more carefully in the new version.

Line 66 - 67 – The fundamental requirements for the evolution of multicellularity are not a bottleneck and reproductive specialisation. Multicellularity is a very broad term, encompassing many facultative forms that only transiently exist as multicellular groups (e.g. *Dictyostelium,* many ciliates and algal species) that do not go through strong bottlenecks or have lifetime reproductive specialization. Even if we are talking about the requirements for obligate, complex multicellularity, I would argue that the fundamental requirement is clonality between the cells (where relatedness = 1), and in fact these are the only lineages where we ever see the evolution of obligate multicellularity. Reproductive specialization is an effect, not a cause, of some multicellular groups.

In this paper we are specifically focused on an evolutionary transition in individually, i.e. the

evolution of Darwinian Individuals. We discuss *Dictyostelium* and other cooperative groups in light of our findings, as examples of groups whose ecology has prevented the evolution of paradigmatic (full-blown) individuality.

As to the second point, we agree on the importance of the bottleneck (clonality). We argue here and in our previous work that reproductive specialization is essential for the evolution of individuality. As discussed above, individuality is a requirement for group adaptation, and we experimentally show in our previous work that reproductive specialization was the essential "difference-maker" for the ETI; in the treatment in which SM germ ("cheats") were suppressed, increased group fitness was a result of increased cell fitness. Reproductive specialization, initially by mutation, resulted in an evolutionary transition. Once adaptations can be selected at the group-level, reproductive specialization can come under developmental control (this occurred in one of the lineages).

Perhaps our disagreement on this point is related to your question above regarding the germ cells here not being a true germ line because they arise (initially) by mutation. If this is the case, then we are in agreement: indeed, the development control of the life cycle can only evolve *as a result of* selection at the group level.

In any case, we have listened to your concerns, and the new introduction takes a more moderate stance on this issue. We hope it is more appealing.

Line 71 – 72 – I would argue that this question is posed incorrectly. The higher-level doesn't have to 'constrain' the lower-level in true multicellular individuals – as they are formed from clonal cells. Therefore this question is only relevant to non-clonal multicellular groups, where relatedness is normally lower and therefore conflict is present between the different cells. ETIs have never occurred with non-clonal multicellular groups, so this question is misleading.

Conflict arises by mutation, even in clonal groups. Here all groups are clonal. Even so, this sentence has been removed.

Line 119 – 122 – Again, I disagree that 'decoupling of fitness' is necessary for transitions to higher levels of individuality. In fact, much evidence suggests that it is in fact the alignment of fitness interests between the cell-level and the group-level (resulting from r = 1) that allows the transition to occur.

Discussed above. We agree so-called 'decoupling' of fitness at the short-term timescale of within-group selection can be interpreted as an alignment of fitness over the long-term: the timescale of the life cycle. The life cycle in the Non-Mixed treatment allows selection to act on the longer timescale, and the apparent short-term reduction of fitness in these groups can be viewed as alignment of fitness in the long-term. We understand that these concepts are difficult for the non-specialist reader, and have taken care in the new version to be more coherent.

Line 352 – 358 – I would argue that these life cycles are a consequence, rather than a cause of the transition to obligate multicellularity organisms.

With respect, we disagree. Multicellularity – at least in paradigmatic form – requires the evolution of Darwinian characters at the level of the collective. One of these characters is reproduction. Lifecycles are how groups acquire the capacity to leave offspring groups. In our view, shared increasingly by others, for example Gestel & Tarnita, the evolution of lifecycles is intimately connected to the evolutionary transition to multicellularity.