THE ROYAL SOCIETY
PUBLISHING

PROCEEDINGS B

Recurrent mutualism breakdown events in a legume rhizobia metapopulation

Kelsey A. Gano-Cohen, Camille E. Wendlandt, Khadija Al Moussawi, Peter J. Stokes, Kenjiro W. Quides, Alexandra J. Weisberg, Jeff. H. Chang and Joel L. Sachs

Article citation details
Proc. R. Soc. B **287**: 20192549.
http://dx.doi.org/10.1098/rspb.2019.2549

Review timeline

Original submission:	13 August 2019
1st revised submission:	31 October 2019
2nd revised submission:	2 January 2020
Final acceptance:	3 January 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2019-1868.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Acceptable

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
No

Is it clear?
N/A

Is it adequate?
N/A

Do you have any ethical concerns with this paper?
No

Comments to the Author
This study measured symbiotic phenotypes of 85 Bradyrhizobium strains associated with the legume Acmispon strigosus and then mapped phenotypes on a phylogenetic tree for two non-symbiotic loci (glnII, recA) to infer patterns of character evolution. The main conclusion was that two key aspects of symbiotic competence (the ability to induce nodule development on A. strigosus, or in cases where nodules did develop, the ability to improve plant growth through nitrogen fixation), had both evolved recurrently (on the order of six times for each trait), in widely separated parts of the bacterial tree (Fig. 2).

Although the general result (recurrent loss of effective symbiosis) may be correct, concerns exist about some of the analyses used to support that conclusion.

The 85 isolates span a large number of Bradyrhizobium lineages (Fig. 2). Most or all of these lineages are also associated with other legume host taxa in a variety of locations. For example, haplotypes "200_G232_R232" and "157_G223_R213" may have a closer relationship to strain CNPSo 4018 (from the legume Indigofera colutea in Australia) than to any of the other Acmispon strigosus isolates in this sample. The 21 isolates with the "G03_R01" haplotype are very close to strain WSM4349 (from Acmispon glaber in Riverside Co., CA). What this means is that Fig. 2 is not a tree showing diversification of strains descended from a single original Acmispon strigosus symbiont. Instead, an unknown number of the 85 isolates represent lineages that evolved in association with other legumes and subsequently colonized this host.
Those strains may have had symbiotic traits that were not adapted to Acmispon strigosus at the time they first arrived, so would have been initially classified as ineffective or partially effective. It is likely that they later acquired the capacity to better fix nitrogen for this legume if they persisted as A. strigosus symbionts. These events are not visible in current reconstruction of trait evolution. So the current snapshot of symbiotic phenotypes mainly depicts scattered losses of symbiotic effectiveness. The analysis lacks the power to detect evolutionary gain of symbiotic effectiveness on A. strigosus because original founder genotypes within each of these Bradyrhizobium lineages are not available to be studied.
Moreover, the timing of colonization is not known, and the ongoing prevalence of migration between symbiont populations associated with Acmispon strigosus and other host taxa is also uncertain. Thus, the symbiotic ineffectiveness of certain strains may not be something that "evolved recurrently" in response to ecological interaction with A. strigosus legumes.
The authors address this issue in the Discussion (p. 14). Their first point is that "none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of

A. strigosus (42, 47) or on different Acmispon species (56)." However, it should be noted that ref. 56 only tested a single ineffective strain (which may be "2_G14_R14" of the current study). Thus, it is hard to say whether A. strigosus-ineffective strains are typically bad for other legumes as well. The authors also argued that "if the ineffective strains were specialized on hosts other than A. strigosus, they should be derived on divergent branches of the phylogeny (especially on the SI tree) rather than being closest relatives with beneficial strains that do appear adapted to A. strigosus". However, some strains that are symbiotically effective on A. strigosus are also effective on other Acmispon species (ref. 56), so this argument is not valid. If strains transfer freely across various legume host taxa, a strain that originally interacted well with two hosts could lose its effectiveness on one of them. There is no reason to expect that this would only happen for strains that are "derived on divergent branches of the phylogeny."

Tests for phylogenetic signal (Table 1) and ancestral state reconstruction (Fig. 2) require a tree-like pattern of evolution for the concatenated glnII + recA data. A large number of the strains do seem to show congruent relationships across the two loci. But BLAST searches and network analysis (using Splitstree4) suggest that there are cases where strains have different relationships for glnII vs. recA (e.g., the large group of isolates with the "G03_R01" haplotype). So it appears that gene transfer events have altered the genealogy of one locus relative to the other. Consequently, no single tree can represent all relationships. Inferences about character evolution on the "tree" (Fig. 2) must be taken with a grain of salt. However, with such a high preponderance of effective isolates (73 of the 85 strains), the scattered distribution of ineffective strains across the tips of the phylogeny seems like a high-probability outcome regardless of the underlying tree topology. So perhaps phylogenetic uncertainty doesn't matter much in the end.

other comments:

lines 106-7: "Nodule and root surface isolates were previously sequenced for glnII and recA, on the Bradyrhizobium chromosome (CHR), and nodZ and nolL, on the symbiosis island (SI) (44, 45)." -- The SI also resides on the chromosome in Bradyrhizobium, so this description is a little misleading. The first two loci are designated as "CHR" but since the nodZ and nolL loci also occur on the chromosome, it would be better to come up with an alternative name, for example "non-SI" or "HK" ("housekeeping").

line 112: "gaps were treated as missing data" -- the nature of gaps needs to be described better. Were there true insertion/deletion polymorphisms involving entire codons, or are the gaps actual missing data just due sequencing limitations?

lines 113-115: "Phylogenetic trees were reconstructed with MrBayes 3.1.2 (55) using 5 × 106 generations, a heating temperature of 0.01, a 'burnin' of 10,000 trees". Tree sampling frequency was stated in line 117 to be every 500 generations, so the total number of trees sampled across 5 million generations would have been only 10,000. So if 10,000 trees were discarded as burnin, there would have been no data retained for phylogenetic inferences. Thus, it appears that one of these values has been stated incorrectly (either the burnin or the total MCMC run length or the sampling frequency).

lines 217-219: "We...report the mean ± standard error of K and average p-values calculated across 200 trees to account for phylogenetic uncertainty." -- how were the 200 trees selected? Were these just a random sample from the trees retained in the MrBayes analysis?

lines 244-245: "six isolates (#'s 2, 155, 186, 187, 200, CW1) 245 from five collection sites were categorized as ineffective". The term "loss of nitrogen fixation" is used to describe these ineffective isolates in the Results (e.g., lines 286, 287, 291). The first sentence of the Discussion also refers to "losses of nitrogen fixation".   Even the title of the paper refers to "Evolutionary losses of nitrogen fixation..."   However, the claim that these strains have lost the capacity for nitrogen fixation goes beyond what the data show. It is entirely possible that these strains retain the capacity to fix some amount of nitrogen (at least under certain conditions).

"Ineffective" is a straightforward descriptive term that accurately describes the phenotype measured for these strains (lack of significant increase in host growth, lack of significant reduction in 15N). This is not necessarily equivalent to the loss of capacity to fix nitrogen. If it had been observed, for example, that these strains had a deletion in one of the nitrogenase structural genes, that would clearly cause the loss of all capacity for N fixation. However, the observation that a nodulating strain has not significantly increased host growth (or significantly lowered 15N percent) is not the same thing as demonstrating that the capacity for N fixation has been lost. Acetylene reduction assays for nitrogenase activity would be helpful in resolving this issue; any positive signal, no matter how small, would disprove the claim that ineffective strains have lost the capacity for N fixation.

lines 263-265: "We were unable to PCR amplify SI loci on isolates that failed to nodulate A. strigosus (nodZ, nolL; Table S1), consistent with degradation or absence of the symbiosis island (44). " -- since nodulation is possible without a nodZ or nolL gene, and since high sequence divergence of nodZ or nolL could result in inability to amplify them even if present, more evidence would be needed to infer degradation or absence of the SI.

lines 304-307: "Host growth response and mean individual nodule mass were also positively correlated (□ = 0.4953, P < 0.0001, n=79), suggesting that A. strigosus hosts preferentially reward more nodule resources to beneficial Bradyrhizobium isolates relative to less-beneficial strains" -- In a no-choice (single-strain) inoculation experiment, the host cannot be said to be preferentially rewarding more beneficial partner strains because each plant's partners are all identical. The fact that mean nodule mass is higher for larger plants would also be consistent with the scenario that the optimal investment by plants in resources for N fixation increases as plant size increases. A positive correlation of host growth and mean individual nodule mass might be expected even in the absence of any system of preferential rewarding of more beneficial partners.

Fig. 2: two isolates have blue phenotype labels in the outer circle, a label that does not match any of the possibilities listed in the figure legend. Is this a mistake? Another isolate has a checkmark in its outer circle phenotype label, which seems like an unintended mark.

Review form: Reviewer 2

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Acceptable

General interest: Is the paper of sufficient general interest?
Acceptable

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
This paper presents the results of a survey of the effectiveness rhizobium bacteria interacting with Acmispon plants across California. The authors have sampled rhizobia from the soil, nodule surfaces, and nodules of Acmispon strigosus plants across 10 sites in CA ranging widely in ecological conditions. Most of the sites are in S California, but two from Central CA are also included, for a total of 85 isolates (side note: how can you have 85 isolates from 10 sites but only 7 isolates/site on average as line 122 says?). The authors found that on a reference plant line, 6 did not nodulate, of the 79 that did, 6 were classified as ineffective because they didn't both significantly increase host growth (10 plants for each strain) AND significantly decrease the N isotope ratio compared to uninoculated plants (all plants were lightly fertilized with 15N to detect N fixation). The authors then conduct ancestral trait reconstruction on a phylogeny of the strains constructed from the chromosome and the symbioses island (they use the consensus tree). The main finding is that loss of effectiveness is recurrent (four independent origins in the phylogeny), which the authors argue points to the instability of the symbiosis.

My main comment is that to me the data in fact is better read as pointing to the *stability* of the mutualistic symbiosis. After all, 6/79 strains being effective on a reference plant line under reference conditions I'd argue shows that by and large, the mutualism is stable. The authors allude to this in the discussion (paragraph starting l 325) when they also point out that the ineffective strains tend to be on the tips of the phylogeny, but nonetheless choose to emphasize instability aspect. I agree with most of this paragraph and its representation of the findings in the paper, but it directly contradicts the last sentence of the abstract. So, I think that emphasis needs to be toned down at least.

Another relevant point to this is the finding that the mean nodule mass and plant growth response were positively correlated, which also points to the stability of mutualism, through partner fidelity feedback at least. It also happens to directly contradict the story in the same group's recent Ecology Letters paper (ref 40 here), where they found no correlation between growth response and in plants fitness, but argue, through a negative correlation with genotype frequencies that there is a fitness conflict even in a single inoculum setting. Checking the excel files for the two papers, it appears that the set of strains is at least partially non-overlapping (the Ecology Letters paper has only 6 populations), which might account for the difference. In any case, I'd argue that nodule biomass is a more direct measure of rhizobium fitness than a snapshot of genotype frequencies in a sample. I find it odd that the authors did not discuss this big difference with their major paper this year and instead cite earlier papers. I appreciate that data are messy and sometimes you get contradictory results, but it seems to me that those results should be acknowledged.

The authors also dismiss alternative hypotheses for the apparent ineffectiveness like GxG interactions a little too hastily. First, they only consider between-species GxG interaction, and not within-species variation — they dismiss this by saying “none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of A. strigosus (42, 47)” but the citations provided here, from the same research group, have only ever used a single ineffective strain, #2. This is a fairly gross misrepresentation of their own work and the authors simply cannot argue that GxG interactions are absent based on those citations. (Ironically, strain #2 in this experiment actually does appear to have enhanced plant growth from the excel sheet in the SI, though to be sure, we’re talking about 79*2 comparisons, so it’s very possible this is a false positive.) Within species GxG would also disarm their second argument, since it wouldn’t necessarily predict the ineffective strains to be outliers phylogenetically.

Minor comments:

Which population is the “reference” plant line derived? The line identifier is given but giving the population it came from would also be useful. Is there a pattern with distance of where the strain is collected from this reference plant population? That would be interesting to know in view of potential local adaptation and isolation by distance.

I don’t understand the sentence on l 299-301: why is this inconsistent with white nodules being ineffective? — after re-reading the methods, I understand this, but it is easy to miss the line in the methods where bulk soil strains are only collected from white or yellow nodules.

Decision letter (RSPB-2019-1868.R0)

03-Oct-2019

Dear Professor Sachs:

I am writing to inform you that your manuscript RSPB-2019-1868 entitled "Recurrent evolutionary losses of nitrogen fixation in a Bradyrhizobium metapopulation" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Professor Loeske Kruuk
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Two expert reviewers have now provided detailed comments on the manuscript. Both express fundamental concerns about whether the data presented are conclusive enough to motivate the broad conclusions of the authors. Most important are the concerns about the interpretation of the phylogenetic analyses, claims of the loss of nitrogen fixation when this was not actually measured, and relating these results to other recent findings in the same system. It will be important for the authors to do some work to consider the broader conceptual points raised by both reviewers about the nature of the inferences that can be made from a tree that likely represents only a subset of lineages and also what the standard is for concluding that mutualism is "unstable" when reconstructing ancestral states on a tree.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
This study measured symbiotic phenotypes of 85 Bradyrhizobium strains associated with the legume Acmispon strigosus and then mapped phenotypes on a phylogenetic tree for two non-symbiotic loci (glnII, recA) to infer patterns of character evolution. The main conclusion was that two key aspects of symbiotic competence (the ability to induce nodule development on A. strigosus, or in cases where nodules did develop, the ability to improve plant growth through nitrogen fixation), had both evolved recurrently (on the order of six times for each trait), in widely separated parts of the bacterial tree (Fig. 2).

Although the general result (recurrent loss of effective symbiosis) may be correct, concerns exist about some of the analyses used to support that conclusion.

The 85 isolates span a large number of Bradyrhizobium lineages (Fig. 2). Most or all of these lineages are also associated with other legume host taxa in a variety of locations. For example, haplotypes "200_G232_R232" and "157_G223_R213" may have a closer relationship to strain CNPSo 4018 (from the legume Indigofera colutea in Australia) than to any of the other Acmispon strigosus isolates in this sample. The 21 isolates with the "G03_R01" haplotype are very close to strain WSM4349 (from Acmispon glaber in Riverside Co., CA). What this means is that Fig. 2 is not a tree showing diversification of strains descended from a single original Acmispon strigosus symbiont. Instead, an unknown number of the 85 isolates represent lineages that evolved in association with other legumes and subsequently colonized this host.

Those strains may have had symbiotic traits that were not adapted to Acmispon strigosus at the time they first arrived, so would have been initially classified as ineffective or partially effective. It is likely that they later acquired the capacity to better fix nitrogen for this legume if they persisted as A. strigosus symbionts. These events are not visible in current reconstruction of trait evolution. So the current snapshot of symbiotic phenotypes mainly depicts scattered losses of symbiotic effectiveness. The analysis lacks the power to detect evolutionary gain of symbiotic effectiveness on A. strigosus because original founder genotypes within each of these Bradyrhizobium lineages are not available to be studied.

Moreover, the timing of colonization is not known, and the ongoing prevalence of migration between symbiont populations associated with Acmispon strigosus and other host taxa is also uncertain. Thus, the symbiotic ineffectiveness of certain strains may not be something that "evolved recurrently" in response to ecological interaction with A. strigosus legumes.

The authors address this issue in the Discussion (p. 14). Their first point is that "none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of A. strigosus (42, 47) or on different Acmispon species (56)." However, it should be noted that ref. 56 only tested a single ineffective strain (which may be "2_G14_R14" of the current study). Thus, it is hard to say whether A. strigosus-ineffective strains are typically bad for other legumes as well. The authors also argued that "if the ineffective strains were specialized on hosts other than A. strigosus, they should be derived on divergent branches of the phylogeny (especially on the SI tree) rather than being closest relatives with beneficial strains that do appear adapted to A. strigosus". However, some strains that are symbiotically effective on A. strigosus are also effective on other Acmispon species (ref. 56), so this argument is not valid. If strains transfer freely across various legume host taxa, a strain that originally interacted well with two hosts could lose its effectiveness on one of them. There is no reason to expect that this would only happen for strains that are "derived on divergent branches of the phylogeny."

Tests for phylogenetic signal (Table 1) and ancestral state reconstruction (Fig. 2) require a tree-like pattern of evolution for the concatenated glnII + recA data. A large number of the strains do seem to show congruent relationships across the two loci. But BLAST searches and network analysis (using Splitstree4) suggest that there are cases where strains have different relationships for glnII vs. recA (e.g., the large group of isolates with the "G03_R01" haplotype). So it appears that gene transfer events have altered the genealogy of one locus relative to the other. Consequently, no single tree can represent all relationships. Inferences about character evolution on the "tree" (Fig. 2) must be taken with a grain of salt. However, with such a high preponderance of effective isolates (73 of the 85 strains), the scattered distribution of ineffective strains across the tips of the phylogeny seems like a high-probability outcome regardless of the underlying tree topology. So perhaps phylogenetic uncertainty doesn't matter much in the end.

other comments:

lines 106-7: "Nodule and root surface isolates were previously sequenced for glnII and recA, on the Bradyrhizobium chromosome (CHR), and nodZ and nolL, on the symbiosis island (SI) (44, 45)." -- The SI also resides on the chromosome in Bradyrhizobium, so this description is a little misleading. The first two loci are designated as "CHR" but since the nodZ and nolL loci also occur on the chromosome, it would be better to come up with an alternative name, for example "non-SI" or "HK" ("housekeeping").

line 112: "gaps were treated as missing data" -- the nature of gaps needs to be described better. Were there true insertion/deletion polymorphisms involving entire codons, or are the gaps actual missing data just due sequencing limitations?

lines 113-115: "Phylogenetic trees were reconstructed with MrBayes 3.1.2 (55) using 5 × 106 generations, a heating temperature of 0.01, a ‘burnin’ of 10,000 trees". Tree sampling frequency was stated in line 117 to be every 500 generations, so the total number of trees sampled across 5 million generations would have been only 10,000. So if 10,000 trees were discarded as burnin, there would have been no data retained for phylogenetic inferences. Thus, it appears that one of these values has been stated incorrectly (either the burnin or the total MCMC run length or the sampling frequency).

lines 217-219: "We...report the mean ± standard error of K and average p-values calculated across 200 trees to account for phylogenetic uncertainty." -- how were the 200 trees selected? Were these just a random sample from the trees retained in the MrBayes analysis?

lines 244-245: "six isolates (#'s 2, 155, 186, 187, 200, CW1) 245 from five collection sites were categorized as ineffective". The term "loss of nitrogen fixation" is used to describe these ineffective isolates in the Results (e.g., lines 286, 287, 291). The first sentence of the Discussion also refers to "losses of nitrogen fixation". Even the title of the paper refers to "Evolutionary losses of nitrogen fixation..." However, the claim that these strains have lost the capacity for nitrogen fixation goes beyond what the data show. It is entirely possible that these strains retain the capacity to fix some amount of nitrogen (at least under certain conditions).
"Ineffective" is a straightforward descriptive term that accurately describes the phenotype measured for these strains (lack of significant increase in host growth, lack of significant reduction in 15N). This is not necessarily equivalent to the loss of capacity to fix nitrogen. If it had been observed, for example, that these strains had a deletion in one of the nitrogenase structural genes, that would clearly cause the loss of all capacity for N fixation. However, the observation that a nodulating strain has not significantly increased host growth (or significantly lowered 15N percent) is not the same thing as demonstrating that the capacity for N fixation has been lost. Acetylene reduction assays for nitrogenase activity would be helpful in resolving this issue; any positive signal, no matter how small, would disprove the claim that ineffective strains have lost the capacity for N fixation.

lines 263-265: "We were unable to PCR amplify SI loci on isolates that failed to nodulate A. strigosus (nodZ, nolL; Table S1), consistent with degradation or absence of the symbiosis island (44). " -- since nodulation is possible without a nodZ or nolL gene, and since high sequence divergence of nodZ or nolL could result in inability to amplify them even if present, more evidence would be needed to infer degradation or absence of the SI.

lines 304-307: "Host growth response and mean individual nodule mass were also positively correlated (□ = 0.4953, P < 0.0001, n=79), suggesting that A. strigosus hosts preferentially reward more nodule resources to beneficial Bradyrhizobium isolates relative to less-beneficial strains" -- In a no-choice (single-strain) inoculation experiment, the host cannot be said to be preferentially rewarding more beneficial partner strains because each plant's partners are all identical. The fact that mean nodule mass is higher for larger plants would also be consistent with the scenario that the optimal investment by plants in resources for N fixation increases as plant size increases. A positive correlation of host growth and mean individual nodule mass might be expected even in the absence of any system of preferential rewarding of more beneficial partners.

Fig. 2: two isolates have blue phenotype labels in the outer circle, a label that does not match any of the possibilities listed in the figure legend. Is this a mistake? Another isolate has a checkmark in its outer circle phenotype label, which seems like an unintended mark.

Referee: 2

Comments to the Author(s)
This paper presents the results of a survey of the effectiveness rhizobium bacteria interacting with Acmispon plants across California. The authors have sampled rhizobia from the soil, nodule surfaces, and nodules of Acmispon strigosus plants across 10 sites in CA ranging widely in ecological conditions. Most of the sites are in S California, but two from Central CA are also included, for a total of 85 isolates (side note: how can you have 85 isolates from 10 sites but only 7 isolates/site on average as line 122 says?). The authors found that on a reference plant line, 6 did not nodulate, of the 79 that did, 6 were classified as ineffective because they didn't both

significantly increase host growth (10 plants for each strain) AND significantly decrease the N isotope ratio compared to uninoculated plants (all plants were lightly fertilized with 15N to detect N fixation). The authors then conduct ancestral trait reconstruction on a phylogeny of the strains constructed from the chromosome and the symbioses island (they use the consensus tree). The main finding is that loss of effectiveness is recurrent (four independent origins in the phylogeny), which the authors argue points to the instability of the symbiosis.

My main comment is that to me the data in fact is better read as pointing to the *stability* of the mutualistic symbiosis. After all, 6/79 strains being effective on a reference plant line under reference conditions I'd argue shows that by and large, the mutualism is stable. The authors allude to this in the discussion (paragraph starting l 325) when they also point out that the ineffective strains tend to be on the tips of the phylogeny, but nonetheless choose to emphasize instability aspect. I agree with most of this paragraph and its representation of the findings in the paper, but it directly contradicts the last sentence of the abstract. So, I think that emphasis needs to be toned down at least.

Another relevant point to this is the finding that the mean nodule mass and plant growth response were positively correlated, which also points to the stability of mutualism, through partner fidelity feedback at least. It also happens to directly contradict the story in the same group's recent Ecology Letters paper (ref 40 here), where they found no correlation between growth response and in plants fitness, but argue, through a negative correlation with genotype frequencies that there is a fitness conflict even in a single inoculum setting. Checking the excel files for the two papers, it appears that the set of strains is at least partially non-overlapping (the Ecology Letters paper has only 6 populations), which might account for the difference. In any case, I'd argue that nodule biomass is a more direct measure of rhizobium fitness than a snapshot of genotype frequencies in a sample. I find it odd that the authors did not discuss this big difference with their major paper this year and instead cite earlier papers. I appreciate that data are messy and sometimes you get contradictory results, but it seems to me that those results should be acknowledged.

The authors also dismiss alternative hypotheses for the apparent ineffectiveness like GxG interactions a little too hastily. First, they only consider between-species GxG interaction, and not within-species variation — they dismiss this by saying "none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of A. strigosus (42, 47)" but the citations provided here, from the same research group, have only ever used a single ineffective strain, #2. This is a fairly gross misrepresentation of their own work and the authors simply cannot argue that GxG interactions are absent based on those citations. (Ironically, strain #2 in this experiment actually does appear to have enhanced plant growth from the excel sheet in the SI, though to be sure, we're talking about 79*2 comparisons, so it's very possible this is a false positive.) Within species GxG would also disarm their second argument, since it wouldn't necessarily predict the ineffective strains to be outliers phylogenetically.

Minor comments:

Which population is the "reference" plant line derived? The line identifier is given but giving the population it came from would also be useful. Is there a pattern with distance of where the strain is collected from this reference plant population? That would be interesting to know in view of potential local adaptation and isolation by distance.

I don't understand the sentence on l 299-301: why is this inconsistent with white nodules being ineffective? — after re-reading the methods, I understand this, but it is easy to miss the line in the methods where bulk soil strains are only collected from white or yellow nodules.

Author's Response to Decision Letter for (RSPB-2019-1868.R0)

See Appendix A.

RSPB-2019-2549.R0

Review form: Reviewer 2

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
This paper presents data on the effectiveness of Bradyrhizobium strains isolated from across California on an Acmispon host plant, and asks how widespread ineffective strains are and where they are in the phylogeny of strains. The patterns reported here are interesting and will be

useful to researchers working on the evolution of cooperative symbioses, but there were some issues with the interpretation.

The first version of the manuscript had a heavy slant towards claiming the mutualism is unstable, but I felt that this was unjustified since more than 90% of all strains (including those isolated from the soil) were in fact effective. In this revision, the authors toned down the emphasis on instability of the mutualism, and rather talk about "recurrent mutualism breakdown." This is a better way to talk about it, but calling the ineffectiveness of a given strain with the test host genotype under the particular experimental conditions "mutualism breakdown" still seems to make a lot of assumptions. I think that the authors should at least acknowledge the possibility that mutualism breakdown might be a very "local" outcome and these strains might actually be effective under different environmental contexts. This is related to the other reviewer's comment about loss of nitrogen fixation: at best they have shown loss of fixation under these particular conditions.

The authors also added some new analyses in response to the other reviewer's comments, which I found helpful and generally increased the strength of the paper. The authors argue that some of the ineffective strains indeed where more related on some of the loci to rhizobia from other hosts, but none in all four, and some effective strains also were more closely related to strains from other hosts. I think the point would be stronger if they could offer a statistical test for whether ineffectiveness makes it more likely to be more closely related to another host's strain.

The authors also improved the discussion, which now more explicitly points out that the current data shows a positive correlation between host and rhizobium fitness, in contrast to their other recent results. If I wanted to nitpick, I would say that the "mutation selection balance" they discuss can hardly be called a "null hypothesis" as it encompasses very non-trivial mechanisms like partner choice or sanctions. They could also point out that the non-correlation found in their other paper is based on a indirect measure of aggregate fitness (strain frequency) whereas the current study uses a more direct measure of in plant fitness, which could explain the difference.

Overall, I think the paper has improved, and does present interesting data. I think with some minor revisions would make it better still.

Review form: Reviewer 3

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Excellent

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
N/A

Is it clear?
Yes

Is it adequate?
No

Do you have any ethical concerns with this paper?
No

Comments to the Author
Gano-Cohen et al. collected rhizobia from multiple wild populations of Acmispon strigosus. They isolated from nodules, bulk soil, and the root surface, which allowed them to potentially isolate non-nodulating rhizobia. They then sequenced several housekeeping and symbiosis genes, which they used to construct phylogenetic trees. Greenhouse experiments were performed to identify isolates that either did not fix nitrogen or did not form nodules on A. strigosus, then the nodulation and fixation phenotypes were mapped onto the trees to see how many times they had been lost, and how deep in the phylogeny.

The essence of the results is:
1. There were several non-nodulating, and several nodulating, but non-fixing isolates recovered.
2. These fix- or nod- strains were (mostly) closely related to fix+, nod+ strains, indicating that the loss of beneficial symbiosis was recent and occurred multiple times. Also, based on the phylogeny, it did not seem that the losses were just adaptation to other host species, although this was not experimentally tested.

I think this is an interesting paper, the approach is reasonable, and in general is well-executed, although the manuscript does need proofreading.

I have a few suggestions that I think will improve the paper:

- Discuss the importance of the sampling scheme to the results. That is, how might deliberately sampling white/yellow nodules, and choosing a certain number of wild isolated vs. trap isolated vs. root surface isolated strains affect the number of loses of nod and fix inferred? Is it possible that most of the rhizobia outside of the nodules are actually either non-nodulating or ineffective, which might lead to an interpretation of frequent *gain* of nod or fix as well as loss?

- Show both the SI and CHR trees constructed for just the 85 focal isolates with branch lengths, maybe as additional supplemental figures. It seems like the existing trees just show topology, but this doesn't give a sense of how close or distant the isolates are from each other

- Are there any non-rhizobial species that are closely related to A. strigosus Bradyrhizobium strains? For example, Agrobacterium and Rhizobium are intermingled phylogenetically.

Also, some minor issues:

- Please carefully check numbering of supplemental tables and figures, some of them are incorrect, and it was a little bit confusing.

- Were trees constructed from nucleotide or protein sequences? Was this the same for the trees constructed from non-A. strigosus isolates too?

- Citation 46 and citation 21 are the same paper

- General proofreading is needed

- line 248: Is my understanding correct that there were two isolates that increased plant growth without forming visible nodules? Seems interesting.

- Please make sure all supplemental tables have legends. S10 does not.

- Table S10: summary for glnII, isolate 200 is incorrect

- Line 297: "...in no case was an ineffective strain more closely related to an isolate from another host species when considering the data from all four loci...". This needs some clarification--do the authors mean that no ineffective strain was more closely related to another host's strain across all the loci? across the majority of loci? when considering distances from the concatenated sequences?

- On Fig. S3, it would be helpful to mark the strains considered to be ineffective

- Fig. S5: some branches are black and some are gray. Is this meaningful?

Decision letter (RSPB-2019-2549.R0)

16-Dec-2019

Dear Professor Sachs

I am pleased to inform you that your manuscript RSPB-2019-2549 entitled "Recurrent mutualism breakdown events in a legume rhizobia metapopulation" has been accepted for publication in Proceedings B.

The referees and Associate Editor have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

To revise your manuscript, log into https://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Before uploading your revised files please make sure that you have:
1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".

2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

5) Data accessibility section and data citation
It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository.

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:
- DNA sequences: Genbank accessions F234391-F234402
- Phylogenetic data: TreeBASE accession number S9123
- Final DNA sequence assembly uploaded as online supplemental material
- Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311

NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link http://datadryad.org/submit?journalID=RSPB&manu=(Document not available) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit https://royalsociety.org/journals/authors/author-guidelines/.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Yours sincerely,
Professor Loeske Kruuk
Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member
Comments to Author:
The paper was re-read by reviewer 2 and was also evaluated by one new reviewer. Both thought that the authors did a nice job responding to reviewer criticisms, including additional analyses of relatedness to other strains that might have adapted to non-Acmispon hosts and improved discussion. Both reviewers of the resubmitted manuscript have additional comments for authors to address in a revision.

In addition to the reviewer comments, I would like the authors address the degree to which the ancestral state reconstructions are sensitive to the initial representation of strains in the tree, and thus how much the directionality of the inference depends on the initial choice of strains. The 85-strain sample is heavily weighted towards strains isolated from nodules, which are most often nodulating and effective – about 3/4 of the strains in the study. To what degree does this influence the direction of ancestral state reconstruction and thus the direction of the transitions inferred? In other words, if the study were done instead using 15 nodule isolates and 62 soil/rhizosphere isolates with many non-mutualistic strains, would the ancestral states be inferred as non-nodulating and thus the conclusion be that the transitions toward mutualism are repeatable and frequent? I am not certain whether there is a statistical way to ask this question. I believe this is a separate question from whether the mutualistic and non-mutualistic strains are intermixed, which seems clear, and also separate from the initial reviewer's concern about alternate hosts. It seems clear that these states are labile over short timescales, which is compelling either way – my question is whether the directionality of these repeated changes depends on the starting sample.

Reviewer(s)' Comments to Author:
Referee: 3

Comments to the Author(s).
Gano-Cohen et al. collected rhizobia from multiple wild populations of Acmispon strigosus. They isolated from nodules, bulk soil, and the root surface, which allowed them to potentially isolate non-nodulating rhizobia. They then sequenced several housekeeping and symbiosis genes, which they used to construct phylogenetic trees. Greenhouse experiments were performed to identify isolates that either did not fix nitrogen or did not form nodules on A. strigosus, then the nodulation and fixation phenotypes were mapped onto the trees to see how many times they had been lost, and how deep in the phylogeny.

The essence of the results is:
1. There were several non-nodulating, and several nodulating, but non-fixing isolates recovered.
2. These fix- or nod- strains were (mostly) closely related to fix+, nod+ strains, indicating that the loss of beneficial symbiosis was recent and occurred multiple times. Also, based on the phylogeny, it did not seem that the losses were just adaptation to other host species, although this was not experimentally tested.

I think this is an interesting paper, the approach is reasonable, and in general is well-executed, although the manuscript does need proofreading.

I have a few suggestions that I think will improve the paper:

- Discuss the importance of the sampling scheme to the results. That is, how might deliberately sampling white/yellow nodules, and choosing a certain number of wild isolated vs. trap isolated vs. root surface isolated strains affect the number of loses of nod and fix inferred? Is it possible that most of the rhizobia outside of the nodules are actually either non-nodulating or ineffective, which might lead to an interpretation of frequent *gain* of nod or fix as well as loss?

- Show both the SI and CHR trees constructed for just the 85 focal isolates with branch lengths, maybe as additional supplemental figures. It seems like the existing trees just show topology, but this doesn't give a sense of how close or distant the isolates are from each other

- Are there any non-rhizobial species that are closely related to A. strigosus Bradyrhizobium strains? For example, Agrobacterium and Rhizobium are intermingled phylogenetically.

Also, some minor issues:

- Please carefully check numbering of supplemental tables and figures, some of them are incorrect, and it was a little bit confusing.

- Were trees constructed from nucleotide or protein sequences? Was this the same for the trees constructed from non-A. strigosus isolates too?

- Citation 46 and citation 21 are the same paper

- General proofreading is needed

- line 248: Is my understanding correct that there were two isolates that increased plant growth without forming visible nodules? Seems interesting.

- Please make sure all supplemental tables have legends. S10 does not.

- Table S10: summary for glnII, isolate 200 is incorrect

- Line 297: "...in no case was an ineffective strain more closely related to an isolate from another host species when considering the data from all four loci...". This needs some clarification--do the authors mean that no ineffective strain was more closely related to another host's strain across all the loci? across the majority of loci? when considering distances from the concatenated sequences?

- On Fig. S3, it would be helpful to mark the strains considered to be ineffective

- Fig. S5: some branches are black and some are gray. Is this meaningful?

Referee: 2

Comments to the Author(s).
This paper presents data on the effectiveness of Bradyrhizobium strains isolated from across California on an Acmispon host plant, and asks how widespread ineffective strains are and where they are in the phylogeny of strains. The patterns reported here are interesting and will be

useful to researchers working on the evolution of cooperative symbioses, but there were some issues with the interpretation.

The first version of the manuscript had a heavy slant towards claiming the mutualism is unstable, but I felt that this was unjustified since more than 90% of all strains (including those isolated from the soil) were in fact effective. In this revision, the authors toned down the emphasis on instability of the mutualism, and rather talk about “recurrent mutualism breakdown.” This is a better way to talk about it, but calling the ineffectiveness of a given strain with the test host genotype under the particular experimental conditions “mutualism breakdown” still seems to make a lot of assumptions. I think that the authors should at least acknowledge the possibility that mutualism breakdown might be a very “local” outcome and these strains might actually be effective under different environmental contexts. This is related to the other reviewer’s comment about loss of nitrogen fixation: at best they have shown loss of fixation under these particular conditions.

The authors also added some new analyses in response to the other reviewer’s comments, which I found helpful and generally increased the strength of the paper. The authors argue that some of the ineffective strains indeed where more related on some of the loci to rhizobia from other hosts, but none in all four, and some effective strains also were more closely related to strains from other hosts. I think the point would be stronger if they could offer a statistical test for whether ineffectiveness makes it more likely to be more closely related to another host’s strain.

The authors also improved the discussion, which now more explicitly points out that the current data shows a positive correlation between host and rhizobium fitness, in contrast to their other recent results. If I wanted to nitpick, I would say that the “mutation selection balance” they discuss can hardly be called a “null hypothesis” as it encompasses very non-trivial mechanisms like partner choice or sanctions. They could also point out that the non-correlation found in their other paper is based on a indirect measure of aggregate fitness (strain frequency) whereas the current study uses a more direct measure of in plant fitness, which could explain the difference.

Overall, I think the paper has improved, and does present interesting data. I think with some minor revisions would make it better still.

Author's Response to Decision Letter for (RSPB-2019-2549.R0)

See Appendix B.

Decision letter (RSPB-2019-2549.R1)

03-Jan-2020

Dear Professor Sachs

I am pleased to inform you that your manuscript entitled "Recurrent mutualism breakdown events in a legume rhizobia metapopulation" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Paper charges
An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit https://royalsociety.org/journals/ethics-policies/media-embargo for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

Dear Professor Sachs:

I am writing to inform you that your manuscript RSPB-2019-1868 entitled "Recurrent evolutionary losses of nitrogen fixation in a Bradyrhizobium metapopulation" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A ‘response to referees’ document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,

Professor Loeske Kruuk
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Two expert reviewers have now provided detailed comments on the manuscript. Both express fundamental concerns about whether the data presented are conclusive enough to motivate the broad conclusions of the authors. Most important are the concerns about the interpretation of the phylogenetic analyses, claims of the loss of nitrogen fixation when this was not actually measured, and relating these results to other recent findings in the same system. It will be important for the authors to do some work to consider the broader conceptual points raised by both reviewers about the nature of the inferences that can be made from a tree that likely represents only a subset of lineages and also what the standard is for concluding that mutualism is “unstable” when reconstructing ancestral states on a tree.
We appreciate these concerns and we have extensively edited our manuscript to address the criticisms. In particular we have added new phylogenetic analyses that consider the genetic relationships between our isolates and rhizobia isolated from other host species, as detailed below.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
This study measured symbiotic phenotypes of 85 Bradyrhizobium strains associated with the legume Acmispon strigosus and then mapped phenotypes on a phylogenetic tree for two non-symbiotic loci (glnII, recA) to infer patterns of character evolution. The main conclusion was that two key aspects of symbiotic competence (the ability to induce nodule development on A. strigosus, or in cases where nodules did develop, the ability to improve plant growth through nitrogen fixation), had both evolved recurrently (on the order of six times for each trait), in widely separated parts of the bacterial tree (Fig. 2).

Although the general result (recurrent loss of effective symbiosis) may be correct, concerns exist about some of the analyses used to support that conclusion.

The 85 isolates span a large number of Bradyrhizobium lineages (Fig. 2). Most or all of these lineages are also associated with other legume host taxa in a variety of locations. For example, haplotypes "200_G232_R232" and "157_G223_R213" may have a closer relationship to strain CNPSo 4018 (from the legume Indigofera colutea in Australia) than to any of the other Acmispon strigosus isolates in this sample. The 21 isolates with the "G03_R01" haplotype are very close to strain WSM4349 (from Acmispon glaber in Riverside Co., CA). What this means is that Fig. 2 is not a tree showing diversification of strains descended from a single original Acmispon strigosus symbiont. Instead, an unknown number of the 85 isolates represent lineages that evolved in association with other legumes and subsequently colonized this host.

We appreciate the concerns of this reviewer, in particular the idea that the ineffective Bradyrhizobium (on A. strigosus) could be adapted to other host species. We now provide extensive additional data to test this hypothesis. Specifically, we reconstructed gene trees for each of the four loci in our study that also incorporate all the related homologous sequences that are available on GenBank (NCBI).

With these data we tested whether i) Bradyrhizobium isolated from other host species are intermixed on the phylogeny -- suggesting that the bacterial populations we are studying also can interact with other hosts, and if so, ii) whether ineffective strains are more likely more closely related to isolates from other hosts (compared to beneficial isolate) – suggesting that rhizobial adaptation to other host plants drives the loss of effectiveness on A. strigosus.

We found that the reviewer is correct in that some Bradyrhizobium isolates from other legumes species are intermixed on the different gene trees (though this is not true for one of the symbiosis loci). However, inconsistent with adaptation to these other hosts, there were no cases where any of our ineffective strains were more closely related to isolates from other species across the whole genome. We discuss these new data extensively in our edited materials and methods (lines 119-125), results (289-297), and discussion 348-355). We appreciate this reviewer input.

Those strains may have had symbiotic traits that were not adapted to Acmispon strigosus at the time they first arrived, so would have been initially classified as ineffective or partially effective. It is likely that they later acquired the capacity to better fix nitrogen for this legume if they persisted as A. strigosus symbionts. These events are not visible in current reconstruction of trait evolution. So the current snapshot of symbiotic phenotypes mainly depicts scattered losses of symbiotic effectiveness. The analysis lacks the power to detect evolutionary gain of symbiotic effectiveness on A. strigosus because original founder genotypes within each of these Bradyrhizobium lineages are not available to be studied.

Moreover, the timing of colonization is not known, and the ongoing prevalence of migration between symbiont populations associated with Acmispon strigosus and other host taxa is also uncertain. Thus, the symbiotic ineffectiveness of certain strains may not be something that "evolved recurrently" in response to ecological interaction with A. strigosus legumes.

We agree with the reviewer's concerns that that we are only focusing on interactions between Bradyrhizobium and A. strigosus hosts. The reviewer is also correct that we did not seek to reconstruct the gain of capacity to associate with A. strigosus on the bacterial tree - nor the timing of such events. However, as stated above, we have now added a test of the degree to which these bacteria are interacting with other host plants. This new dataset adds a nice dimension to the analysis.

The authors address this issue in the Discussion (p. 14). Their first point is that "none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of A. strigosus (42, 47) or on different Acmispon species (56)." However, it should be noted that ref. 56 only tested a single ineffective strain (which may be "2_G14_R14" of the current study). Thus, it is hard to say whether A. strigosus-ineffective strains are typically bad for other legumes as well.

We understand this criticism, and in light of the new data we have deleted these phrases.

The authors also argued that "if the ineffective strains were specialized on hosts other than A. strigosus, they should be derived on divergent branches of the phylogeny (especially on the SI tree) rather than being closest relatives with beneficial strains that do appear adapted to A. strigosus". However, some strains that are symbiotically effective on A. strigosus are also effective on other Acmispon species (ref. 56), so this argument is not valid. If strains transfer freely across various legume host taxa, a strain that originally interacted well with two hosts could lose its effectiveness on one of them. There is no reason to expect that this would only happen for strains that are "derived on divergent branches of the phylogeny."

Given the new data that we have added, we have deleted this argument.

Tests for phylogenetic signal (Table 1) and ancestral state reconstruction (Fig. 2) require a tree-like pattern of evolution for the concatenated glnII + recA data. A large number of the strains do seem to show congruent relationships across the two loci. But BLAST searches and network analysis (using Splitstree4) suggest that there are cases where strains have different relationships for glnII vs. recA (e.g., the large group of isolates with the "G03_R01" haplotype). So it appears that gene transfer events have altered the genealogy of one locus relative to the other. Consequently, no single tree can represent all relationships. Inferences about character evolution on the "tree" (Fig. 2) must be taken with a grain of salt. However, with such a high preponderance of effective isolates (73 of the 85 strains), the scattered distribution of ineffective strains across the tips of the phylogeny seems like a high-probability outcome regardless of the underlying tree topology. So perhaps phylogenetic uncertainty doesn't matter much in the end.

We agree with the concerns of the reviewer that 'no single tree can represent all relationships.' This is why we have used the Bayesian approach of analyzing data on a set trees, rather than individual reconstructions. In this case, for instance, analyses of phylogenetic signal were tested using randomizations on a set of 200 trees (see lines 217-220).

other comments:

lines 106-7: "Nodule and root surface isolates were previously sequenced for glnII and recA, on the Bradyrhizobium chromosome (CHR), and nodZ and nolL, on the symbiosis island (SI) (44, 45)." -- The SI also resides on the chromosome in Bradyrhizobium, so this description is a little misleading. The first two loci are designated as "CHR" but since the nodZ and nolL loci also occur on the chromosome, it would be better to come up with an alternative name, for example "non-SI" or "HK" ("housekeeping"). *We have added a bit of important explanation here. The reviewer is correct that the SI is often integrated on the chromosome. But it need not be and sometimes exists as a plasmid (we added a citation). Given* *that i) the SI need not be integrated, and ii) we have used the CHR and SI terms in several other studies, we would prefer to keep these labels.*

line 112: "gaps were treated as missing data" -- the nature of gaps needs to be described better. Were there true insertion/deletion polymorphisms involving entire codons, or are the gaps actual missing data just due sequencing limitations?
We reviewed the specific data here and found that the missing data involved only a single base pair that we were not able to resolve after multiple attempts. Since this base did not contain any parsimony informative characters we decided to delete this sentence completely.

lines 113-115: "Phylogenetic trees were reconstructed with MrBayes 3.1.2 (55) using 5 × 106 generations, a heating temperature of 0.01, a 'burnin' of 10,000 trees". Tree sampling frequency was stated in line 117 to be every 500 generations, so the total number of trees sampled across 5 million generations would have been only 10,000. So if 10,000 trees were discarded as burnin, there would have been no data retained for phylogenetic inferences. Thus, it appears that one of these values has been stated incorrectly (either the burnin or the total MCMC run length or the sampling frequency).
We have clarified that the 'burnin' period represents the first 10,000 trees (leaving 4,990,000 remaining which were sampled with a frequency of 1/500). Thus we sampled 9980 post burnin trees.

lines 217-219: "We...report the mean ± standard error of K and average p-values calculated across 200 trees to account for phylogenetic uncertainty." -- how were the 200 trees selected? Were these just a random sample from the trees retained in the MrBayes analysis?
We have clarified that we used 200 randomly selected post-burnin trees.

lines 244-245: "six isolates (#'s 2, 155, 186, 187, 200, CW1) 245 from five collection sites were categorized as ineffective". The term "loss of nitrogen fixation" is used to describe these ineffective isolates in the Results (e.g., lines 286, 287, 291). The first sentence of the Discussion also refers to "losses of nitrogen fixation**". Even the title of the paper refers to "Evolutionary losses of nitrogen fixation..." However, the claim that these strains have lost the capacity for nitrogen fixation goes beyond what the data show. It is entirely possible that these strains retain the capacity to fix some amount of nitrogen (at least under certain conditions).**

In all these cases we now add clarification that the loss is specific to the Bradyrhizobium association with A. strigosus. We have also changed the title of the manuscript so that we no longer use the phrase "evolutionary losses of nitrogen fixation". The new title is: "Recurrent mutualism breakdown events in a legume rhizobia metapopulation". This more general term relates to the evolution of uncooperative mutants from mutualist ancestors, which is the process that we were seeking to test in this host-symbiont setting.

"Ineffective" is a straightforward descriptive term that accurately describes the phenotype measured for these strains (lack of significant increase in host growth, lack of significant reduction in 15N). This is not necessarily equivalent to the loss of capacity to fix nitrogen. If it had been observed, for example, that these strains had a deletion in one of the nitrogenase structural genes, that would clearly cause the loss of all capacity for N fixation. However, the observation that a nodulating strain has not significantly increased host growth (or significantly lowered 15N percent) is not the same thing as demonstrating that the capacity for N fixation has been lost. Acetylene reduction assays for nitrogenase activity would be helpful in resolving this issue; any positive signal, no matter how small, would disprove the claim that ineffective strains have lost the capacity for N fixation.

We appreciate these criticisms and agree that the terminology used is critical. In our revised manuscript, we have been careful to state that the lack of nitrogen fixation can only be determined in the context of this specific host, A. strigosus. However, we note that we DID measure nitrogen fixation via isotopic analysis, which is considered by most researchers to be superior in accuracy to the Acetylene reduction assay. See the classic paper by Minchin et al. 1983 J. Exp. Bot. 34:641).

lines 263-265: "We were unable to PCR amplify SI loci on isolates that failed to nodulate A. strigosus (nodZ, nolL; Table S1), consistent with degradation or absence of the symbiosis island (44). " -- since nodulation is possible without a nodZ or nolL gene, and since high sequence divergence of nodZ or nolL could result in inability to amplify them even if present, more evidence would be needed to infer degradation or absence of the SI.
We understand the reviewer's criticism and we have restated the sentence to more clearly reference the published dataset of Hollowell et al. 2016 (Ref #44). In that paper we also used these two genes in addition to several other symbiosis genes and we found that they were either all present or all absent, consistent with the degradation or loss of the whole island. See lines 264-266.

lines 304-307: "Host growth response and mean individual nodule mass were also positively correlated ($\rho = 0.4953$, $P < 0.0001$, n=79), suggesting that A. strigosus hosts preferentially reward more nodule resources to beneficial Bradyrhizobium isolates relative to less-beneficial strains" -- In a no-choice (single-strain) inoculation experiment, the host cannot be said to be preferentially rewarding more beneficial partner strains because each plant's partners are all identical. The fact that mean nodule mass is higher for larger plants would also be consistent with the scenario that the optimal investment by plants in resources for N fixation increases as plant size increases. A positive correlation of host growth and mean individual nodule mass might be expected even in the absence of any system of preferential rewarding of more beneficial partners.
We agree with the reviewer's criticism and have made this statement broader as suggested above (see lines 309-315).

Fig. 2: two isolates have blue phenotype labels in the outer circle, a label that does not match any of the possibilities listed in the figure legend. Is this a mistake? Another isolate has a checkmark in its outer circle phenotype label, which seems like an unintended mark.
We corrected the two blue phenotype labels, which was an error. We cannot find the checkmark in this figure.

Referee: 2

Comments to the Author(s)
This paper presents the results of a survey of the effectiveness rhizobium bacteria interacting with Acmispon plants across California. The authors have sampled rhizobia from the soil, nodule surfaces, and nodules of Acmispon strigosus plants across 10 sites in CA ranging widely in ecological conditions. Most of the sites are in S California, but two from Central CA are also included, for a total of 85 isolates (side note: how can you have 85 isolates from 10 sites but only 7 isolates/site on average as line 122 says?).
We have corrected this error.

The authors found that on a reference plant line, 6 did not nodulate, of the 79 that did, 6 were classified as ineffective because they didn't both significantly increase host growth (10 plants for each strain) AND significantly decrease the N isotope ratio compared to uninoculated plants (all plants were lightly fertilized with 15N to detect N fixation). The authors then conduct ancestral trait reconstruction on a phylogeny of the strains constructed from the chromosome and the symbioses island (they use the consensus tree). The main finding is that loss of effectiveness is recurrent (four independent origins in the phylogeny), which the authors argue points to the instability of the symbiosis.

My main comment is that to me the data in fact is better read as pointing to the *stability* of the mutualistic symbiosis. After all, 6/79 strains being effective on a reference plant line under reference conditions I'd argue shows that by and large, the mutualism is stable. The authors allude to this in the discussion (paragraph starting l 325) when they also point out that the ineffective strains tend to be on the tips of the phylogeny, but nonetheless choose to emphasize instability aspect. I agree with most of this paragraph and its representation of the findings in the paper, but it directly contradicts the last sentence of the abstract. So, I think that emphasis needs to be toned down at least.
We have reduced the emphasis on the overall instability of the mutualism and instead make the arguments more specific to the detected transitions and their drivers.

Another relevant point to this is the finding that the mean nodule mass and plant growth response were positively correlated, which also points to the stability of mutualism, through partner fidelity feedback at least. It also happens to directly contradict the story in the same group's recent Ecology Letters paper (ref 40 here), where they found no correlation between growth response and in plants fitness, but argue, through a negative correlation with genotype frequencies that there is a fitness conflict even in a single inoculum setting. Checking the excel files for the two papers, it appears that the set of strains is at least partially non-overlapping (the Ecology Letters paper has only 6 populations), which might account for the difference. In any case, I'd argue that nodule biomass is a more direct measure of rhizobium fitness than a snapshot of genotype frequencies in a sample. I find it odd that the authors did not discuss this big difference with their major paper this year and instead cite earlier papers. I appreciate that data are messy and sometimes you get contradictory results, but it seems to me that those results should be acknowledged.
We have now completely rewritten this section to deal with these criticisms. We clearly explain the differences between results in our study and the two other studies that found evidence of cheating.

The authors also dismiss alternative hypotheses for the apparent ineffectiveness like GxG interactions a little too hastily. First, they only consider between-species GxG interaction, and not within-species variation — they dismiss this by saying "none of the ineffective strains have ever been found to fix nitrogen, whether inoculated on divergent lines of A. strigosus (42, 47)" but the citations provided here, from the same research group, have only ever used a single ineffective strain, #2. This is a fairly gross misrepresentation of their own work and the authors simply cannot argue that GxG interactions are absent based on those citations. (Ironically, strain #2 in this experiment actually does appear to have enhanced plant growth from the excel sheet in the SI, though to be sure, we're talking about 79*2 comparisons, so it's very possible this is a false positive.) Within species GxG would also disarm their second argument, since it wouldn't necessarily predict the ineffective strains to be outliers phylogenetically.
We have rewritten this section to deal with these criticisms. We now more precisely outline the previous data, and also explain the new data that we used to test for GxG interactions (i.e., conditional ineffectiveness).

Minor comments:

Which population is the "reference" plant line derived? The line identifier is given but giving the population it came from would also be useful. Is there a pattern with distance of where the strain is collected from this reference plant population? That would be interesting to know in view of potential local adaptation and isolation by distance.
We added the information that the test line was from our Claremont population.

I don't understand the sentence on l 299-301: why is this inconsistent with white nodules being ineffective? — after re-reading the methods, I understand this, but it is easy to miss the line in the methods where bulk soil strains are only collected from white or yellow nodules.*We added additional explanation into that section.*

Appendix B

Dear Editor & Associate Editor at Proceedings B,

I am pleased to submit a revision of the manuscript RSPB-2019-2549 entitled "Recurrent mutualism breakdown events in a legume rhizobia metapopulation". I have pasted the comments of the Associate Editor and reviewers below and in each case I explain how I have responded in our revised manuscript. Thank you very much for your time and attention on our manuscript.

Cheers,

Joel L. Sachs

Associate Editor
Board Member
Comments to Author:
The paper was re-read by reviewer 2 and was also evaluated by one new reviewer. Both thought that the authors did a nice job responding to reviewer criticisms, including additional analyses of relatedness to other strains that might have adapted to non-Acmispon hosts and improved discussion. Both reviewers of the resubmitted manuscript have additional comments for authors to address in a revision.

Author Response: thank you for positive comments and further attention to our manuscript. We appreciate the constructive criticism and feel that this has led to an overall improvement of the quality of the manuscript.

In addition to the reviewer comments, I would like the authors address the degree to which the ancestral state reconstructions are sensitive to the initial representation of strains in the tree, and thus how much the directionality of the inference depends on the initial choice of strains. The 85-strain sample is heavily weighted towards strains isolated from nodules, which are most often nodulating and effective – about 3/4 of the strains in the study. To what degree does this influence the direction of ancestral state reconstruction and thus the direction of the transitions inferred? In other words, if the study were done instead using 15 nodule isolates and 62 soil/rhizosphere isolates with many non-mutualistic strains, would the ancestral states be inferred as non-nodulating and thus the conclusion be that the transitions toward mutualism are repeatable and frequent? I am not certain whether there is a statistical way to ask this question. I believe this is a separate question from whether the mutualistic and non-mutualistic strains are intermixed, which seems clear, and also separate from the initial reviewer's concern about alternate hosts. It seems clear that these states are labile over short timescales, which is compelling either way – my question is whether the directionality of these repeated changes depends on the starting sample.

Author Response: Thank you for this suggestion. We have added a new portion of the discussion to explore this possibility, i.e., that gains of symbiosis islands could lead to the origins of ineffective strains. Using data from a previous publication (Hollowell et al. 2016 *Microbial Ecology* 71:700) we suggest that as many as three of the ineffective isolates in the

current study might have evolved from non-nodulating ancestors, though this still leaves three independent origins of ineffective strains from beneficial ancestors. See lines 410-424.

Reviewer(s)' Comments to Author:

Referee: 3

Comments to the Author(s).
Gano-Cohen et al. collected rhizobia from multiple wild populations of Acmispon strigosus. They isolated from nodules, bulk soil, and the root surface, which allowed them to potentially isolate non-nodulating rhizobia. They then sequenced several housekeeping and symbiosis genes, which they used to construct phylogenetic trees. Greenhouse experiments were performed to identify isolates that either did not fix nitrogen or did not form nodules on A. strigosus, then the nodulation and fixation phenotypes were mapped onto the trees to see how many times they had been lost, and how deep in the phylogeny.

The essence of the results is:
1. There were several non-nodulating, and several nodulating, but non-fixing isolates recovered.
2. These fix- or nod- strains were (mostly) closely related to fix+, nod+ strains, indicating that the loss of beneficial symbiosis was recent and occurred multiple times. Also, based on the phylogeny, it did not seem that the losses were just adaptation to other host species, although this was not experimentally tested.

I think this is an interesting paper, the approach is reasonable, and in general is well-executed, although the manuscript does need proofreading.

Author Response: thank you for positive comments. We have proofread the manuscript to remove any remaining errors.

I have a few suggestions that I think will improve the paper:

- Discuss the importance of the sampling scheme to the results. That is, how might deliberately sampling white/yellow nodules, and choosing a certain number of wild isolated vs. trap isolated vs. root surface isolated strains affect the number of loses of nod and fix inferred? Is it possible that most of the rhizobia outside of the nodules are actually either non-nodulating or ineffective, which might lead to an interpretation of frequent *gain* of nod or fix as well as loss?

Author Response: We have added a section in the discussion that focuses on the relative effects of sampling and the interpretations of both gains and losses of symbiosis traits, See lines 420-424.

- Show both the SI and CHR trees constructed for just the 85 focal isolates with branch lengths, maybe as additional supplemental figures. It seems like the existing trees just show topology, but this doesn't give a sense of how close or distant the isolates are from each other

We have added two additional supplemental figures with Bayesian phylograms of the sampled isolates for the SI and CHR loci (Figures S2, S5).

- Are there any non-rhizobial species that are closely related to A. strigosus Bradyrhizobium strains? For example, Agrobacterium and Rhizobium are intermingled phylogenetically.

No, we only uncovered other *Bradyrhizobium* that were closely related to the isolates sampled in our study.

Also, some minor issues:

- Please carefully check numbering of supplemental tables and figures, some of them are incorrect, and it was a little bit confusing.
This has been corrected.

- Were trees constructed from nucleotide or protein sequences? Was this the same for the trees constructed from non-A. strigosus isolates too?
We used DNA data in all cases. We have added the word 'nucleotide' in lines 110 and 122.

- Citation 46 and citation 21 are the same paper
This has been corrected.

- General proofreading is needed
We have proofread the manuscript to remove any remaining errors.

- line 248: Is my understanding correct that there were two isolates that increased plant growth without forming visible nodules? Seems interesting.
Yes, but the effects are modest, so we did not spend a lot of time discussing this.

- Please make sure all supplemental tables have legends. S10 does not.
This has been corrected.

- Table S10: summary for glnII, isolate 200 is incorrect
This has been corrected.

- Line 297: "...in no case was an ineffective strain more closely related to an isolate from another host species when considering the data from all four loci...". This needs some clarification--do the authors mean that no ineffective strain was more closely related to another host's strain across all the loci? across the majority of loci? when considering distances from the concatenated sequences?
This has been changed to focus on the SI loci that are most informative. It now says "Genetic distance matrices of CHR loci showed that ineffective strains were sometimes more closely related to isolates from other legume species than to beneficial strains from *A. strigosus*.

However, this was never the case for the SI loci or when taking all four loci into account (table S12). (Lines 321-324)

- On Fig. S3, it would be helpful to mark the strains considered to be ineffective
This has been corrected.

- Fig. S5: some branches are black and some are gray. Is this meaningful?
Figure legends have been changed on this and similar trees (now S7,S8,S10,S11).

Referee: 2

Comments to the Author(s).
This paper presents data on the effectiveness of Bradyrhizobium strains isolated from across California on an Acmispon host plant, and asks how widespread ineffective strains are and where they are in the phylogeny of strains. The patterns reported here are interesting and will be useful to researchers working on the evolution of cooperative symbioses, but there were some issues with the interpretation.

The first version of the manuscript had a heavy slant towards claiming the mutualism is unstable, but I felt that this was unjustified since more than 90% of all strains (including those isolated from the soil) were in fact effective. In this revision, the authors toned down the emphasis on instability of the mutualism, and rather talk about "recurrent mutualism breakdown." This is a better way to talk about it, but calling the ineffectiveness of a given strain with the test host genotype under the particular experimental conditions "mutualism breakdown" still seems to make a lot of assumptions. I think that the authors should at least acknowledge the possibility that mutualism breakdown might be a very "local" outcome and these strains might actually be effective under different environmental contexts. This is related to the other reviewer's comment about loss of nitrogen fixation: at best they have shown loss of fixation under these particular conditions.
Author Response: thank you for positive comments. We have changed some wording to highlight the fact that these are local transitions (see Lines 370, 373).

The authors also added some new analyses in response to the other reviewer's comments, which I found helpful and generally increased the strength of the paper. The authors argue that some of the ineffective strains indeed where more related on some of the loci to rhizobia from other hosts, but none in all four, and some effective strains also were more closely related to strains from other hosts. I think the point would be stronger if they could offer a statistical test for whether ineffectiveness makes it more likely to be more closely related to another host's strain.
Author Response: We now more clearly highlight our summary analysis of genetic distances that takes all the genetic data into account. See lines 404-410.

The authors also improved the discussion, which now more explicitly points out that the current data shows a positive correlation between host and rhizobium fitness, in contrast to their other recent results. If I wanted to nitpick, I would say that the "mutation selection balance" they discuss can hardly be called a "null hypothesis" as it encompasses very non-trivial mechanisms like partner choice or sanctions.

Author Response: We agree with this, but the literature has repeatedly shown that the mechanisms of partner choice and sanctions are both very robust in this association such that ineffective strains get punished.

They could also point out that the non-correlation found in their other paper is based on a indirect measure of aggregate fitness (strain frequency) whereas the current study uses a more direct measure of in plant fitness, which could explain the difference.
Author Response: We have already explained the different methods of fitness. We do not agree that one is more direct than the other.

Overall, I think the paper has improved, and does present interesting data. I think with some minor revisions would make it better still.