Dear Editor,
We have now revised our manuscript according to the comments of the recommender and of the two reviewers. We want to thank these three persons for their advices, comments and feedback which allowed improving the amplitude, clarity and impact of our analysis.
We hope this new version, which substantially changed, as you will realize, can be recommended by PCI Evolutionary Biology.
Sincerely

Decision

by Nicolas Galtier, 2018-09-19 22:10
Manuscript: https://doi.org/10.1101/362129

Decision on: "Parallel adaptations...", by Koutsovoulos et al.

The two reviewers have expressed relevant and important comments on various aspects of the study. I concur that the current manuscript should be extensively revised in order to reach, and convince, a wider evolutionary biology audience, as expected for a manuscript recommended by PCI.

Reviewer 2 recapitulates the main results of the study and identifies a number of issues requiring clarification, rewriting, and/or re-analysis. This reviewer also suggests that the current title does not optimally reflect the content of the study - I agree.

We changed the title to better reflects the main results and messages of the paper. We added a whole new panel of analyses that now confirms the lack of evidence for recombination and we show that the different host races most likely reflect multiple independent gains and losses of parasitic abilities.

In addition, Reviewer 1 mentions a number of analyses that could be made in order to better characterize the population genomics and molecular evolution of M. incognita, with a focus on its supposed asexuality. I agree that this is a missed opportunity, especially knowing that previous publications on the subject, some by authors of this manuscript, have opened very interesting questions (eg Castagnone-Sereno & Danchin 2014 JEB).
The two reviews are highly complementary and provide a number of clearly expressed recommendations, which I think should greatly help improve the manuscript.

We totally agree and have profoundly reinvestigated the data to indeed characterize the population genomics in the light of the mode of reproduction of M. incognita. This is a whole new part of the revised papers and now a main result.

Additional requirements of the managing board
We ask you to carefully verify that your manuscript complies with the following requirements (indicated in the 'How does it work?' section and in the code of conduct) and to modify your manuscript accordingly:

-Data must be available to readers after recommendation, either in the text or through an open data repository such as Zenodo, Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.
→ All the data generated have already been deposited in GEO
To review GEO accession GSE116847: Go to
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116847
-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) must be available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.
The different scripts and R codes used to process the data are available on GitHub at the following URL:https://github.com/GDKO/gdk_scripts/tree/master/popgenvcf
-Details on experimental procedures must be available to readers in the text or as appendices.
We provided more details on the methods and experimental procedure and produced a detailed supporting supplementary material document.

Reviews

Reviewed by anonymous reviewer, 2018-08-25 00:35
In this manuscript, the authors collected 11 'isolates' of the parasitic root knot nematode Meloidogyne incognita from 6 different crops. These isolates were then assigned to four 'host races' traditionally recognized in the species based on the ability of nematodes to infect particular reference crop species (these reference species differ from the crops the nematodes were collected from). The authors then re-sequenced the genomes of the isolates and inferred SNPs relative to an available reference genome of the species. Two independent clustering approaches based on the SNP data (PCA and a phylogenetic network bases on the subset of SNPs in coding regions) indicate that the 11 isolates form 3 diverged clusters. From these clusters the authors draw the following 3 conclusions:

1. The clusters do not correspond to 'host races' and the use of the term 'host race' should be abandoned for M. incognita. I have no problem with this conclusion - in more 'standard' terminology, host races are polyphyletic and correspond to a phenotype rather than a lineage.

We now statistically tested the significance of the lack of correlation between host races and SNP clusters using a Fisher's exact test. This test checks for independency between the tested variables, based on the co-occurrence counts.

We used the following contingency table of the 12 isolates, plus the Morelos strain (reference genome) whose host race is known (cluster C, Race 3), in clusters and host races:

Cluster / Host Race	R1	R2	R3	R4
A	1	0	0	0
B	0	0	1	1
C	3	2	3	2

The Fisher exact test returns a p-value of 1, indicating no significant association between host races and SNP clusters of isolates.

To evaluate the sensitivity of the Fisher test we also ran the same test on an artificially modified contingency table.
In this modified contingency table, we assigned all the R3 to cluster C and all the R4 to cluster B and left the rest of the table unchanged.

Cluster / Host Race	R1	R2	R3	R4
A	1	0	0	0
B	0	0	0	3
C	3	2	4	0

This modification alone was enough to now obtain a p-value of 0.007, indicating that this method is sensitive in detecting significant association, if present.

These tests reinforced the hypothesis of the non-monophyly of the host races. We think the term race should be abandoned because, it usually implies monophyly whereas other terms like biotype or pathotype are phenotype-based artificial groupings that don't imply a single common ancestral inheritance.

However, the authors also state that these data indicate multiple independent adaptations to different host ranges. In the absence of information of ancestral host ranges, this is clearly an over-interpretation. For example, one could imagine a highly polyphagous ancestral lineage (i.e., with extreme phenotypic plasticity for host plant use), but the ability to infect specific hosts was lost independently in different lineages. The title "Parallel adaptations to different host plants despite clonal reproduction in the world most devastating nematode pest" should therefore be abandoned as it does not reflect the findings of the paper.

We found the hypothesis of the reviewer (a polyphagous ancestor and then specialisation of the ancestral host range towards more restricted host races) interesting and decided to test this. The 'host races' are defined according to the differential capability of infecting just two plant cultivars: tobacco and cotton. Starting from the phylogenetic network produced from the SNP data, we deduced a phylogenetic tree of the 11 isolates that we rooted at the midpoint. We then reconstructed ancestral states, in terms of capacity of infecting cotton and tobacco at each node using maximum likelihood (Fig. S8). According to reviewer 1's hypothesis the ancestral state should be Race 4 (able to parasitize both cotton and tobacco). Then the other races emerged multiple times independently by losing a little bit of the

ancestral polyphagy (e.g. Race3 by losing the ability to parasitize tobacco, etc…). We observed that the most ancestral state shows a 0.5 probability of being able to parasitize Cotton and a 0.5 probability of being able to parasitize Tobacco. Hence the combined probability of being able to parasitize both Cotton and Tobacco (= being Race 4) is only 0.25. This hypothesis is thus not very likely.
The ancestral states reconstruction shows that It is more likely that multiple gains and losses of ability to parasitize Tobacco and / or Cotton occurred during the evolution of the different isolates. The observation that nematodes can switch host race is indicative that adaptation is at least as likely than loss of ability to parasitise. In *M. incognita*, switch of host races during time have been observed in two directions, from Race 2 (compatible with Tobacco but not Cotton) to Race 4 (compatible with both Tobacco and Cotton) and also from Race 4 to Race 2 (R.M.D.G Carneiro, personal observations). Furthermore, it has been shown that M. incognita is able to break down resistance genes in tomato and thus adapt to a new plant cultivar (Castagnone-Sereno et al. 1994; Castagnone-Sereno 2006).
To conclude, in the initial version of the paper we had indeed favored the hypothesis of multiple neo-adaptations vs losses. In this new version, we have rephrased everything, including the title, to better reflect that the current observed 'host races' probably reflect multiple independent gains and losses of abilities to parasitize different plants.

2. The authors state that there is no correlation between the diverged clusters and the geographic origin of the samples. However, according to Fig 5 this is not entirely true as there is some clustering of geographically close samples. Instead of eyeballing whether or not the genetic clusters correspond some grouping of samples from a given country, I suggest the author conduct standard IBD analyses and calculate the % of genetic variance explained by geographic distance (i.e., using pairwise geographic distances between isolates). This % will be small, but represents a more objective evaluation of the effect of geography.

We agree with reviewer 1, association between countries and the PCA clusters can be misleading, particularly because Brazil and the USA are large countries with very distant collected points. An isolation by distance (IBD) is thus more appropriate as it takes into account the geographical distance regardless of the borders of the countries under consideration. We collected as precise as possible geographical coordinates of the fields from which the nematodes had been collected within Brazil and in the rest of the world, then conducted an IBD analysis. Both for Brazil and for the whole worldwide dataset, the IBD analysis showed no evident correlation between Nei's genetic distance and the geographical distance (p-value: 0.718) (Fig. S3).
In addition, it should be kept in mind that, because *M. incognita* is an agricultural pest, most of its current geographical distribution is probably highly impacted by human activities and trades of soils, crops and plants. Hence the lack of correlation between genetic and geographical distance is not very surprising.

3. Similar to point 2, the authors also state that there is no correlation between the clusters and the crop species where the isolates were collected. This is difficult to evaluate given the small number of isolates (n=11) relative to the number of crops (n=6).

Nevertheless, 3 of the 4 isolates from cotton are members of the same cluster, as are the 2 out of 2 isolates from tobacco, so there appears to be some correlation. Again, I suggest that the authors quantify the amount divergence of isolates of the same vs different crop species to show that the amount of variance between crops is not larger than the variance within.
We agree that this must be statistically tested and we used a Fisher exact test on contingency tables between clusters and crop species. This kind of test is appropriate and recommended when the size of the dataset is limited.

For crop species we performed two analysis

- 1. All the isolates for which a crop species was documented regardless the number of occurrences of the crop species (16 isolates on 8 different crops cf. Table S4)

Cluster	Coffee	Cotton	Cucumber	Grape	Soybean	Tobacco	Tomato	Watermel
A	0	0	0	0	1	0	1	0
B	0	1	0	0	1	0	0	0
C	1	4	1	1	0	2	2	1

The Fisher exact test returns a p-value of 0.69, indicating no significant association between crop species and clusters

- 2. All the isolate / crop species associations where host species count was >1 to eliminate potential biases associated to single isolate crop species (12 isolates on 4 different crops cf. Table S4)

Cluster	Cotton	Soybean	Tobacco	Tomato
A	0	1	0	1
B	1	1	0	0
C	4	0	2	2

The Fisher exact test returns a p-value of 0.26, indicating, here again, no significant association between crop species and clusters

Given the results of these two tests we can safely state that, overall, there is no significant association between the SNP-defined clusters of isolates and their crop species of origin (although in some local cases like Tobacco 2/2 isolates are in a same cluster).

Finally, there are parts in the methods that should be clarified. For example, I believe the term ‘isolate’ usually refers to a strain derived from a single female, but here apparently it is a pool of individuals collected at a given location (or even from multiple locations for isolate R3-4, see line 117). How did the authors deal with population variation – was only the major allele considered at each position?

For each isolate, one single female and its associated egg mass were retrieved. To confirm the species (here *M. incognita*), we used esterase isozyme patterns on the female. The single egg mass was used for tomato plant infection and multiplication for a (variable) number of generations. Once enough nematodes were obtained, a pool was collected and we performed the North Carolina Differential Host Test (NCDHT) to characterize the host race, accordingly.The rest of the population was kept for multiplication on tomato plants to produce enough nematodes for sequencing (usually ~1 million individuals pooled together). The nematodes remaining in the soil (and not collected for the pool sequencing) were kept and conserved in liquid nitrogen in the nematode collection at EMBRAPA CENARGEN in Brasilia, Brasil.

Along the same lines: M. incognita is a hybrid species and highly heterozygous. I believe the reference assembly is largely haploid (the two parental genomes are assembled separately), meaning there will be no heterozygous positions for a genome based on a single genotype. These points do not affect the results but should be clarified in the methods.
Yes, exactly, the genome of M. incognita is most likely triploid and of hybrid origin with high average nucleotide divergence (~8%) between the homoeologous regions (Blanc-Mathieu et al. 2017). Because of this high divergence, the assembly was able to correctly separate most of the homoeologous regions. Indeed, the assembled genome size is 183.5 Mb for a total estimated genome size of 189 Mb (+/- 15Mb cf.(Blanc-Mathieu et al. 2017)). Hence, the genome assembly can mostly be considered as effectively haploid. However, in this same paper, it was estimated that ~17Mb of the M. incognita genome had a coverage twice higher than the rest of the genome. We confirmed that the distribution of per-base coverage in the reference genome presents two peaks, with one at twice the coverage of the first one (Fig. S2). The regions with this double coverage most likely represent homoeologous regions with lower divergence that have been collapsed during the assembly. Mapping reads from the two homoeologous haplotypes on one single collapsed haplotype will systematically generate artificial heterozygous SNPs. Supporting this hypothesis, the distribution of per base coverage of heterozygous SNPs (0/1) present a peak at twice the coverage for homozygous (1/1) SNPs (Fig. S1). Hence, we decided to discard the 0/1 heterozygous SNPs from the rest of the analysis because they probably represent artefacts dues to collapsed homoeologous regions. This is now clarified in the new version of the paper.

Minor comments. - In the discussion of divergence from the reference genome for nuclear and mitochondrial sequences (L241 and following), I am surprised the authors are not mentioning lack of recombination – the lack of recombination in mt genomes contributes to their relatively high substitution rates in comparison to the nuclear genome in sexual species. This difference no longer exists in M. incognita.
This is an excellent remark and this absence of recombination could indeed explain why, in contrast to C. elegans, we don't observe the same order of magnitude between the nuclear and mitochondrial DNA divergences in M. incognita. We updated the discussion accordingly. Furthermore, the lack of evidence for recombination in the M. incognita nuclear genome is something we have now tested and confirmed at the population genomics level following comments of Reviewer 2.

-L134, 135: do you mean divergence from the Morelos reference strain? (are these uncorrected p-distances?)
Yes, It is uncorrected p-distance and in terms of divergence from Morelos reference strain, we noticed that reviewer 2 has the same question and we thus clarified this in the text.

-The map in Figure 1 has 13 isolates listed for Brasil, not 11.
Yes, it is because isolate R3-4 is a pool from 3 different states in Brazil, all extracted from cotton fields.

-L108: I dont understand the sentence “Considering the clonal reproduction of this species, the route of adaptation to different hosts is evolutionary important since it is unknown whether it happens independently or a consequence of four ancestral states.”
We agree that this sentence needed clarification and the whole discussion has been totally reformatted and changed according to the new analyzes performed and the main conclusions of the paper.

-‘Host preference’ usually refers to a neurological mechanism of preferring one host over another, not to being able to infect one host but not another. I would chose a different term.
We totally agree with the reviewer, we have no evidence that the nematodes choose their host plants, so ‘host preference’ is not an appropriate term. In the NCDHT, nematodes are indeed forced to infect plants and have no opportunity to choose / prefer one. We thus changed host preference by host compatibility.

Reviewed by anonymous reviewer, 2018-08-25 00:39
This manuscript presents a survey of genetic diversity at the whole genome level in the root-knot nematode Meloidogyne incognita, which is a clonal worldwide plant pathogen with different host races. The main question was to determine whether these races corresponded to distinct genetic clusters. Three main genetic clusters were found but they are neither associated with host races nor geographical origins. This result is interesting with important practical implications and the ms clearly presents this main result.
We reinforced this main message by statistical tests as suggested by Reviewer 1. These tests allowed confirming the absence significant association between host race status, crop species of origin, geographical distribution and SNP-based clusters.

However, I think that the dataset (whole genome sequence of 11 strains + additional published genomes) is clearly under-analysed and that a better knowledge of the system could be reach, also with potential practical implications.
This version should be sufficient for a specific audience specifically interested in the biological model. But to reach a broader audience, additional analyses should be done
We agree with the reviewer comments and have now seized this opportunity to address broader questions such as the evidence or lack of evidence for recombination or the efficiency of selection.

MAJOR COMMENTS

A major assumption is that the species is clonal. However this is not discussed while the data could help to evaluate more precisely the breeding system of the species. Is the species really completely clonal? Is it recent or not?
We totally agree, this assumption has never been backed by population genomics data and in the new version of the paper, we have investigated it, mainly using 4-gametes test and linkage disequilibrium.

So far, the only evidence for mitotic parthenogenesis (or clonality) in Meloidogyne incognita as remained only cytogenetics (Triantaphyllou 1985). No meiosis has ever been observed in this species including in the female gonads at the same location where meiosis is usually observed in the sexual relatives like *Meloidogyne hapla*. The embryos develop from unfertilized eggs that did not undergo any meiotic division (or any change in ploidy). Furthermore, one single larva, can give rise to a whole offspring, alone, in the form of an egg mass without fecondation. Finally, the estimated chromosome numbers indicate a triploidy, which would make 'classica' meiosis really challenging.
For this ensemble of reasons, the offspring is considered clonal, although clonality and absence of presence of meiosis have never been tested at a whole gnome scale.
In this revised version, we used the population genomics date to investigate evidence (or lack thereof) for meiosis.

How recent is clonality / mitotic parthenogenesis?
M. incognita most probably became mitotic parthenogenetic following hybridization between facultative sexual (meiotic parthenogenetic) parent species (Hugall et al. 1999; Blanc-Mathieu et al. 2017). In Blanc-Mathieu et al. 2017, phylogenomics analysis of the genome structure of M. incognita as well as the two close relatives M. javanica and M. arenaria (also mitotic parthenogens) suggests a recent allopolyploidization at the origin of these three species. Indeed, while the different homoeologous copies of the nuclear DNA are highly diverged within a species, the mitochondrial DNA is almost identical between the three species. Hence, the most logical explanation is recent hybridization events between closely related females and more distant males. Furthermore, while most of the gene copies are expected to be rapidly lost after WGD (here allopolyploidization), only very few gene losses between homoeologous regions were observed in the genomes of the parthenogenetic root-knot nematodes. Finally, even if extremely rare, males are sometimes observed under adverse conditions. This ensemble of facts suggest a recent allopolyploid origin.

For example, Fig 4 represents a network and not a fully resolved tree. In a purely clonal species, a perfect tree should be expected. Potential signature of genetic exchanges could be assessed with the current dataset:
Indeed, in the previous version of the paper, Fig 4 was a network. However it became clear that the network structure was caused by 0/1 'heterozygous' SNPs within isolates. These 0/1 heterozygous sites are probably mostly artefacts and not the result of inter-individual heterozygosity within isolates. Indeed, as explained now in the new version, although most of the homoeologous regions have been correctly separated in the *M. incognita* genome assembly (due to their high average sequence divergence), some regions of lower divergence (representing ~20Mb) have been collapsed during the assembly. These

collapsed regions present a per base coverage twice the one observed for the rest of the genome (Fig. S2). One consequence of this is that genome reads of the too homoeologous regions map to a same collapsed region in the assembly and thus a lot of 0/1 SNPs would be artificially called. Consistent with this view, by plotting the base coverage of 0/1 heterozygous SNPs, we realized that they showed a peak at twice the coverage of the 1/1 homozygous SNPs (Fig. S1). Thus, most of these 0/1 SNP probably represent artificial SNP due to collapsed slightly heterozygous regions.
Hence, we discarded heterozygous SNPs from the analysis and conducted the same phylogenetic networks analysis, only keeping the 1/1 homozygous SNPs. The new resulting Fig4 is now a bifurcating tree and not a network anymore.
This shows that the 0/1 SNPs are the ones responsible for creating connections within the branches in the tree. This observation perfectly makes sense with the idea of diverged genome reads representing two homoeologous matching a single collapsed region, in the different isolates..

- Under pure clonality all genomic regions should give the same history. So first did you get the same nuclear and mitochondrial tree?
This is a very good suggestion. We did a mitochondrial phylogeny and, yes, indeed, this perfectly recapitulates the nuclear phylogeny (Fig. S6).

Then, if you split the nuclear dataset into blocks (chromosomes or shorter blocks) do you obtain the same tree for all blocks? This is a simple way to test whether genetic exchanges have occurred or not? -
We totally agree, and indeed, in the absence of recombination, phylogenies made on split datasets (here scaffolds because we don't have chromosome-scale resolution of the genome) should recapitulate the topology obtained with the full dataset. We did this analysis on the 14 biggest scaffolds containing enough informative SNPs. On each occasion, the phylogeny recapitulated, the three previously defined clusters (Fig. S5), with obviously more polytomies within the biggest cluster due to a lower number of informative sites. These results show no evident signs for recombination. We performed more detailed and dedicated tests (see below) to further investigate evidence (or lack thereof) for meiosis.

There are also different methods to test the occurrence of recombination (or gene conversion) that should be applied here (ex: decrease of linkage-disequilibrium with distance, four-gamete test, and more elaborated methods).
This is an excellent suggestion and in this new version of the paper we did extensively search for evidence for recombination. To sum up, we used the following properties: under recombination linkage disequilibrium (LD) between markers should rapidly decrease with the physical distance between these markers. In parallel the proportion of bi-allelic markers that pass the 4-gamete test should rapidly increase with the distance between the markers. Under clonality, in contrast, all the markers should be linked and there should be no clear effect of physical distance on the LD or on the proportion of pairs of markers that pass the 4-gamete test.
By performing the four-gametes test and LD analysis on the whole set of homozygous SNPs, we observed no trend for a decrease in LD with the physical distance between the

markers (Figure 5). Similarly, we observed no increase of the proportion of markers passing the four-gamete test with the distance between bi-allelic markers (Figure 5).
To assess the sensitivity of our analysis, we also performed LD and 4-gametes test on the genome of the obligatory meiotic outcrosser cyst nematode Globodera rostochiensis, for which genome data of different geographical isolates were also available (Eves-van den Akker et al. 2016). Compared to M. incognita, the results obtained on G. rostochiensis are totally opposed and show all the expected signs of meiosis: a rapid decrease of LD with physical distance and a concomitant rapid increase of the proportion of markers passing the 4-gametes test (Fig. 6).
This analysis showed a lack of evidence for recombination in M. incognita at the whole genome level.

- Individual heterozygosity is not given.
The material we have sequenced results from pools of >1 million eggs for each isolate that each derive from the egg mass of a single female.
In principle, individual heterozygosity could be estimated by calculating the variations between individuals within a given isolate, based on the heterozygous '0/1' SNPs. However, as explained above, the 0/1 SNPs are probably mostly artefacts due to collapsed homoeologous regions of low nucleotide divergence. Because the status of the 0/1 heterozygous SNPs is more likely to represent artefacts than variations between individuals, they were discarded from the rest analysis.

Under pure clonality, excess of heterozygosity (Fis <0) is also expected I think this point is crucial to correctly interpret the results. Genetic diversity is only briefly analysed and not very precisely.
Because we discarded heterozygous SNPs from the analysis, for more safety, we could not calculate Fis values.

P6 l133: What is exactly "level of variation"? Does it mean Tajima's pi or another statistics?
This is uncorrected p-distance. We also now have calculated PiN and PiS values (see below).

To allow comparison with other species it would be interesting to compute pi synonymous (or 4-fold): is it of the same order of other clonal or selfing species?
The comparison with selfing and outcrossing nematodes should be particularly relevant. More generally, recent surveys of genetic diversity can be used for comparison, for ex: - Romiguier et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261-263. - Chen, J. et al. 2017. Genetic Diversity and the Efficacy of Purifying Selection across Plant and Animal Species. Mol Biol Evol 34:1417-1428.
To better interpret piS in term of effective population size (Ne), an idea of the mutation rate should be important also.

In addition to piS, the piN/piS ratio should also be computed and given. It gives an idea of the efficiency of selection and is usually rather well correlated with Ne. This could be compared also with previous studies.
The analyses suggested above could be done at the whole species level but also for the three different genetic clusters separately.
This is an excellent suggestion and we have now performed these analyses. We used SNPGenie (Nelson et al. 2015) to calculate nucleotide diversity at synonymous and non synonymous sites (PiS and PiN) as well as the ratio PiN/ PiS in M. incognita and compared these values to those previously obtained for different Caenorhabditis species, from the references cited by reviewer 2.
In the Nature paper, PiS, PiN and PiN/PiS values were calculated for two outcrossing nematodes:
-C.doughertyi (previously named sp.10) : PiS= 0.0493; PiN= 0.0025; PiN/PiS= 0.0513
-C. brenneri: PiS= 0.0322; PiN= 0.0013; PiN/PiS= 0.0396
In the Mol Biol Evol paper, these values have been calculated for on selfer Caenorhabditis species
- C. briggsae: Pi0= 0.0019; Pi4= 0.0074; Pi0/Pi4= 0.258

We obtained the following values for M. incognita:

isolate	piN	piS	piN/piS
R1-2	0.0000856	0.000560	0.153
R1-3	0.0000578	0.000374	0.154
R1-6	0.0000553	0.000353	0.157
R2-1	0.0000463	0.000282	0.164
R2-6	0.0000505	0.000324	0.156
R3-1	0.0000542	0.000348	0.156
R3-2	0.0000669	0.000438	0.153
R3-4	0.0000558	0.000357	0.156
R4-1	0.0000535	0.000341	0.157
R4-3	0.0000545	0.000345	0.158
R4-4	0.0000712	0.000479	0.149

The range of values is PiS= 0.0003 - 0.0006; PiN= 0.00005 - 0.00008; PiN/PiS= 0.149 - 0.164
Hence, compared to the outcrossers, the PiS and PiN values are much lower but the PiN/PiS ratio substantially higher in M. incognita, suggesting purifying selection is less efficient despite overall lower diversity.

Concerning the mutation rate in M. incognita, for the moment there is no reliable estimate and also no information about the separation time between the different isolates. We are actually starting experiments in our lab to accurately estimate the mutation rate but this will take some time.

In addition, Fst should be given to get a more quantitative idea of population structure than just the PCA.
We calculated fixation index (F_{ST}) for the three clusters using vcftools (Danecek et al. 2011) and we obtained the following values:

- Fst for Homozygous SNPs only; upper part mean, lower part weighted

	clusterA	clusterB	clusterC
clusterA	X	0.98088	0.3523
clusterB	0.9904	X	0.52805
clusterC	0.83675	0.84255	X

It would also be interesting to present the distribution of genetic diversity along chromosomes. This could also bring information about potential rare or past recombination events, for example if there is more genetic diversity in telomeric than in centromeric regions.
We agree that this would be interesting but, unfortunately, this cannot be done at a chromosome scale resolution for the moment. The current genome assembly is still fragmentary (>12,000 scaffolds, with a N50 of 38kb) and far from a chromosome-scale resolution of the genome (30-40 chromosomes). This kind of analysis will be possible in the future with more contiguous assembly. It should be noted also that the chromosomes in *M. incognita* and other Meloidogyne are holocentric, and there is no actual centromere (Triantaphyllou 1981).

"M. incognita is particularly versatile and adaptive despite its clonal mode of reproduction": this interpretation implicitly assumed that adaptation to the host is highly multigenic and complex. However, if only a few key genes determine host compatibility, the problem of being clonal is much less important and it maybe not necessary to invoke CNV or epigenetic mechanisms. It would also explain why there is no association between genetic cluster and host race. Because there is no association between genetic cluster and host races it is potentially a good situation to identify genes potentially involved in adaptation to the different hosts.
We agree that adaptive evolution in M. incognita despite its clonal propagation might just be related to a few genes (or changes in some regulatory elements) and that it is not absolutely necessary to invoke mechanisms others than accumulation of SNP and other short-scale variations. Hence, is host compatibility determined by variations in only a few genes? This is an interesting hypothesis and actually was one of the initial goals of our analysis: finding gene variants associated to the host compatibilities.
The 4 host races defined in the NCDHT depend on the compatibility of just two plant cultivars: Tobacco and Cotton. Race 1 is compatible with none of them (0-0), Race 2 is

compatible with Tobacco but not with Cotton (1-0), Race 3 not with Tobacco but Cotton (0-1) and Race 4 is compatible with both (1-1).
We thus tried to identify SNPs in genes, shared by Race 2 and Race 4 that would be associated to compatibility with Tobacco and SNPs in genes, shared by Race 3 and Race 4 that would be associated to compatibility with Cotton. We did not find any such SNP in gene that could be correlated to the ability of parasitizing either tobacco or cotton.
Another reason why we evoke mechanisms others than SNPs and short-scale mutations is that we recently showed associations between multiple convergent losses in gene copy numbers and ability of M. incognita to break down a nematode resistance gene in tomato (Castagnone-Sereno et al. 2019). We could imagine that similar mechanisms might be at play in gains and losses of ability to parasitize different host plants.

The sample size may be too small but more elaborate genomic scan could be done rather than simply searching for specific SNPs. In particular, the genomic location should be taken into account to increase the power of detection.
As explained above, because our genome is currently still fragmentary, analysis of the distribution of the SNPs along chromosomes or chromosomal segments is not yet possible. This will hopefully be the case with a future version of the genome.

MINOR COMMENTS P4 l73: “no clear genetic differences underlying the phenotypic plasticity” this is not well formulated because strictly speaking phenotypic plasticity does not require genetic variation, otherwise we would rather use “local adaptation”
We agree and have reformulated this in the new version.

P5 l96: “in this analysis; no relation” --> “;” should be replaced by “,”
This whole sentence had been removed from the revised version of the paper.

Fig 2: Number of variants per isolate: to what are variants defined? The reference strain?
Yes, it is the number of positions that varies compared to the reference genome (Morelos strain).

Fig4: the scale on the figure is too small and hardly readable. Is it 0.1?
Figure 4 and all the figures have been re-designed for better clarity.

Cited references:
Blanc-Mathieu R, Perfus-Barbeoch L, Aury J-M, Rocha MD, Gouzy J, Sallet E, Martin-Jimenez C, Bailly-Bechet M, Castagnone-Sereno P, Flot J-F, et al. 2017. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. PLOS Genet. 13:e1006777.
Castagnone-Sereno P. 2006. Genetic variability and adaptive evolution in parthenogenetic root-knot nematodes. Heredity 96:282–289.
Castagnone-Sereno P, F. V-M, F. L. 1994. Genetic polymorphism between and within Meloidogyne species detected with RAPD markers. Genome 37:904–909.
Castagnone-Sereno P, Mulet K, Danchin EGJ, Koutsovoulos GD, Karaulic M, Da Rocha M, Bailly-Bechet M, Pratx L, Perfus-Barbeoch L, Abad P. 2019. Gene copy number variations as signatures of adaptive evolution in the parthenogenetic, plant-parasitic nematode *Meloidogyne incognita*. Mol. Ecol.:mec.15095.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Eves-van den Akker S, Laetsch DR, Thorpe P, Lilley CJ, Danchin EGJ, Da Rocha M, Rancurel C, Holroyd NE, Cotton JA, Szitenberg A, et al. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. Genome Biol. [Internet] 17. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0985-1

Hugall A, Stanton J, Moritz C. 1999. Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic Meloidogyne. Mol. Biol. Evol. 16:157–164.

Nelson CW, Moncla LH, Hughes AL. 2015. SNPGenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. Bioinformatics 31:3709–3711.

Triantaphyllou AC. 1981. Oogenesis and the chromosomes of the parthenogenetic Root Knot Nematode Meloidogyne incognita. J. Nematol.:95–104.

Triantaphyllou AC. 1985. Cytogenetics, cytotaxonomy and phylogeny of root-knot nematodes. In: Sasser JN, Carter CC, editors. An advanced treatise on Meloidogyne: Biology and control. Vol. 1. Raleigh, USA: North Carolina State University Graphics. p. 113–126.