Response to reviewer comments

We are very grateful to the reviewers for their careful and thoughtful reviews of the manuscript. We respond to each comment in detail below. We believe that adding requested detail has improved the clarity and quality of the manuscript.

Reviewer comments appear in blue below and responses are in black. (Note that lines numbers in reviewer references are no longer correct because line numbers have changed in the revision.)

**Reviewer 1**

1. Your aim was to describe the diversity of SRs in Cnidaria but you just analysed deeply the SR-B family. Please include the analysis of SR-E and SR-I (the same way done for SR-B) in the Results section, since you found signals of expansion on them too.

We have elected to retain the phylogenetic analysis for the SR-Bs but not add analyses for the other family members identified in the study. We feel that primary focus of the study was on the SR-Bs, because indeed they were the sequences originally identified in previous work to be differentially expressed and therefore likely important in symbiosis. We feel that this study is a first look at SR diversity but that detailed phlylogenetic analysis of all sub-families awaits another study. This is a middle ground between this request by reviewer 1 and the request by reviewer 2, point 38 below, to remove the phylogenetic analysis altogether because it is not detailed enough.

2. The experimental step of the work with fucoidan is fine as a first approach. Thank you.

3. Line 55 – The bootstrap support value from the ancestral node of cnidarians and the rest of bilaterian was 1. It should be written “high support” instead of “moderate support”. We reworded this to make the meaning of the sentence clearer.

4. Line 328 – Please include the information of Figure S1 into Figure 2. Done.

5. Line 346 – There are 18 CD36 for Cnidaria in Figure 2 but you just used 12 in Figure 3. Where are the other six? The missing CD36 are: one from A. elegantissima, one from A. pallida, three from A. digitifera and one from F. scutaria.

The alignment was not created with all the sequences that were identified in our searches. We selected a subset of sequences that we felt best represented the diversity of types. To clarify this, we added “a subset of” to the description of creating the alignment.

6. Line 346 – “SR-B-like” instead of “SRB-like”. Done.

7. Lines 348-367 – Please report the results for SR-E and SR-I too. As detailed in point 1 above, we have chosen not perform alignments on these SR subfamilies.

8. Line 388 – It is not well explained. You said 72 SRCR-domain-containing protein but 32 of them contains only CTLD domains and four of them have only CD36. Also, the number 72 refers to all domains found for Aiptasia (including five in the supplementary file). Please correct this sentence. This was a mistake. Thank you for finding this. We have corrected the error.

9. Lines 156-166 – Why these six species? How many cnidarian species with genome sequenced were available when you assembled the data? You could include the date you downloaded the genomes/transcriptomes here. Also, why you reassembled the data of Aiptasia? Please add the answer in the text.

We included all available cnidarian genomes publically available at the time except *Hydra* which is in a different class from Anthozoa. There is no easy way to include a timestamp on when we worked with the resources. We accessed them multiple times through several years and the resources are dynamic – with annotations constantly changing. We reassembled that *Aiptasia* transcriptome with an assembler that resulted in a better assembly. We have added this on lines 184-185.

10. Line 182 – 1x10-1 is not a high e-value. Delete the high, please. In fact it is a high number – this is a correct statement. High = weak expect value. Low = strong expect value. We therefore left this as is.

11. Lines 199-202 – You used just the CD36 domain to reconstruct the phylogenetic tree (line 205) but the way is written here appears that you used these three domains. This statement does not refer to the phlyogenetic analysis of the SR-Bs but rather to the characterization of the different SR types presented in Figure 2. Therefore we have left this unchanged.

12. Line 204-213 – Why did you not reconstruct the phylogeny of SR-E and SR-I families since there was expansion on these domains in Cnidaria too? Please repeat this analysis (as well as the alignment on Figure S2) for SR-E and SR-I. Also, make the alignments for the three SRs (SR-B, SR-E and SR-I) available (e.g. Dryad repository). Please see points 1 and 7 above.

Lines 204-213 – Please cite the references for the software employed: MAFFT, Geneious, ProtTest, FastTree, SEQBOOT and PhyML. Done.

14. Lines 212-213 – Instead “identical topography” use “identical topology”. Done

**Reviewer 2**

15. The article presents interesting results in the diversity of scavenger receptors (SR) in cnidarians and is generally well written. Results are relevant and properly represented by figures. Thank you

16. Background information in the structure and function of SR are well covered in the introduction, but there is no introduction to the interplay between the immune system and the production of NO.

In the context of cnidarian-dinoflagellate symbiosis, this section is even more important, as bleaching is proposed to be triggered by oxidative stress. A paragraph describing this interplay in the introduction would bridge the gap between the bioinformatic analysis and the laboratory essays, enhancing the text flow.

The use of DAF oxidation as a proxy for NO concentration could be explained in the introduction too.

We respectfully disagree with the above requests. The topic of the study is narrowly focused on SRs in cnidarians and their possible role in symbiosis. The introduction does not specifically cover cnidarian bleaching, oxidative stress and the role of innate immunity in these processes. This is a big topic that would require far more than a paragraph to introduce. Furthermore, the experiments shown in figure 5 use NO and oxidative stress as a read out, a handle on the state of immune system stimulation, and not as evidence of stress or bleaching per se. We feel that citing these studies when describing the methods, in the methods section is the appropriate location for these citations.

17. Some phrases are misplaced and a general review of the manuscript could resolve this issues.

In general, we prefer a writing style where some background or introductory information that is relevant to a method being described is appropriate in the methods. Likewise, we also prefer a style that allows for contextual statements in the methods and/or results that improve both the flow of the manuscript and the clarity of the experiments being proposed. See below for specific comments

The following statements should be moved to the introduction:

18. Lines 255-260: 'We hypothesized that if a symbiont is co-opting host SRs to initiate tolerogenic pathways (such as the TGFβ pathway) that dampen or prevent an immune response, blocking SR-ligand-binding capabilities could induce an immune response upon the addition of lipopolysaccharide (LPS). LPS is a MAMP that has been shown to induce an anemone immune response measured as increased nitric oxide (NO) production (Detournay et al. 2012; Perez & Weis 2006).' We added a sentence at the end of the introduction that lays out this hypothesis, lines 156-158. We also elected to leave the text above in the methods. It provides context on immune stimulation techniques that do not need to be introduced in the introduction.

19. Lines 313-314: 'Vertebrate SR-As are defined by a collagen domain coupled with most proteins containing either an SRCR domain or a CTLD at the C terminus (Bowdish & Gordon 2009)' We elected to leave this in the results as it is directly relevant to the data being presented in that section.

20. Lines 317-319: 'Human SR-Es are defined by the presence of only CTLDs (Zani et al. 2015). The human lectin-like oxidized low-density lipoprotein receptor 1 (LOX1) has an N-terminal cytoplasmic tail, a transmembrane domain and a single C-terminal CTLD (Canton et al. 2013)' We elected to leave this in the results as it is directly relevant to the data being presented in that section.

21. Lines 320-322: 'SR-Is in humans are defined by containing only SRCR domains in various numbers of repeats and are grouped into three classes: CD5, CD6 and CD163' We elected to leave this in the results as it is directly relevant to the data being presented in that section.

The following statements should be moved to the methods:

22. Line 373: '(to test for fucoidan toxicity to the animal)' We removed this clause. There is already mention of this test in the methods on lines 284-285.

23. Lines 376-378: 'A second fucoidan experiment investigated the possible immune-regulation role of an SR in symbiosis maintenance. Symbiotic anemones were treated with increasing concentrations of fucoidan and were subsequently immune-challenged by incubation with LPS.' We elected to leave this sentence in the results as a contextual and clarifying statement.

Minor issues:

24. Lines 158/159: Either indicate it is Aipatsia pallida (as it was done throughout the manuscript) or remove the term 'species', for consistency. (I know this is not the authors' fault, anyway...) We changed everything to *Aiptasia pallida or A. pallida* throughout the manuscript

25. Lines 162-166: There might be some confusion between experimental manipulation and bioinformatic manipulation. Change 'All resources were used without manipulation' for something like 'Published transcriptome/genome assemblies'. We changed wording to improve clarity

26. Line 169: 'primarily from GenBank'. What's the other source? We have added reference to Table S1 that lists the resources

27. Line 170: Most importantly, these sequences are queries for the Blast searches on cnidarian resources. We have added a separate tab labelled blast query sequences in the Supplementary File 1.

28. Line 178-179: not clear what the authors mean by 'Keyword searches using the terms SR, CD36, LMP2, SRCR, and scavenger were also performed.' Was it on the annotation provided by the original articles? How were then obtained? Was it different from the methodology presented? Within the genome online search tools for *N. vectensis* and *A. digitifera* there is a keyword search option that allows searching of the available annotation with keyword identifiers. There is also some GO and KEGG annotation available for the transcriptomes. Where available, this was searched using the keywords listed. Keyword searching is a common term used to describe using keywords to query an annotated database. Please see the Joint Genome Institute website for more information on Keyword searches and KEGG annotation http://genome.jgi.doe.gov/help/genomeSearch.jsf.

29. Lines 190-191: Which databases? We added detail here

30. Line 202. It's actually figure 2. Done

31. Line 260. Reinforce that 'Symbiotic anemones were incubated' Done

32. Line 286: 'a baseline colonization level for symbiotic anemones'. Shouldn't it be 'aposymbiotic anemones'? No, this was a measure of symbiotic animals

33. Line 378-380: Change 'The FSW control-treated anemones had low levels of NO production, a proxy for an immune response, measured as MFI of the NO-specific probe DAF-FM DA in tentacles, in response to incubation in LPS.' to: 'The FSW control-treated anemones had low levels of NO production in response to incubation in LPS.'  Once again, we have elected to keep the text as is to increase the clarity of the experiment and to emphasize that the NO read-out is a measure of an immune response

34. The manuscript clearly defines a meaningful research question, addressed in an original research.
Reporting of the methods can be improved with a clearer reporting of the bioinformatic analysis. blast and pfam searches should be explained separately, with respective queries and algorithm parameters (lines 169 – 197). We have added detail, as requested above, to try to make this section clear.

35. Description of the bioessays shall be clearer if the paragraph on immune response and NO production is included in the introduction. See above point 16 above for our justification for keeping this information brief and in the methods

36. The statistical model should be presented, with more explanation. It is hard to evaluate/understand what was done. Was it used random effects models for the individuals and polyps within individuals? It looks like a nested design with repeated measures is more appropriate for dealing with the lack of independence in lines 298-301.

We have updated the paragraph to reflect these comments (lines 327-342). In particular, we have included an explicit statement of the mixed effects model used for analysis. We also included the data and script used for statistical analysis as a supplement for readers to verify. Our model treats tentacle within individual as the sampling unit, with a random effect for individual anemones. Since we only used one measure of colonization success per tentacle, a random effect for tentacle within anemone would be confounded with residual variation, and would not be estimated correctly. The explicit statement of the statistical model should clarify this. A repeated measures analysis is not warranted here since the same anemones were not used repeatedly – as they were kept in relaxing solution for the confocal microscopy measurements, they had to be discarded after measurements. So measures at different time-points are from different anemones.

37. As the test is based in differences between slopes, the slopes should be presented in the results (at least as supplementary data).

We regret the confusion around the perception about differences in slopes – the test for colonization success with fucoidan treatment concentration and time was not based on slopes per se but on the interaction term between time-point and treatments. When plotted against time as a continuous variable, this is akin to a slope, but since time is used here as a discrete variable (I.e., time-point), the term slope does not really apply. On the other hand, for the NO fluorescence measurements, we used fucoidan concentration as a continuous variable, and fitted a slope. We give this slope, along with the likelihood profile confidence interval in line. We have split the detail of the statistical analysis for each experiment to avoid confusion (see lines 327-342 and lines 354-365).

38. The only retention I have with the results published in this study is with the phylogenetic analysis of CD36 domain containing sequences. Results suggests that CD36-domain containing protein diversification occurred independently in cnidarians and bilaterians. Although this might be interesting, it contributes little to this manuscript. At least in the manner it is currently presented.

Furthermore, 'basal' metazoan sequences other than cnidarians are under-represented and the phylogeny is constructed with varying number of paralogous sequences. In this context, the fact that a ctenophore sequence groups together with sponges instead of cnidarians looks like a long-branch attraction artifact.

My recommendation is to remove it from the current manuscript and treat it in a separate study, where this analysis could be improved with a more comprehensive review of the available genomic resources and by the identification of orthologous sequences.

This request, to either provide a more detailed phylogenetic analysis of SR-Bs or remove these data from the study, goes against the request of Reviewer 1 (point 1) to add additional phylogenetic analyses of other SRs. We argue here for retaining the current phylogeny in the manuscript. Yes it would be ideal to perform a large phylogenetic analysis on SR-Bs with more sequences, including more basal metazoan sequences included in the analysis. However we do feel that this modest analysis adds information to the study. To try to qualify the strength of the phylogenetic results, we have add a sentence on lines 502-504 that reads: “A large analysis including additional sequences from basal metazoans is required to more definitively reveal deep branching patterns of this gene.” As a side note, ctenophores are now considered to be basal to sponges and cnidarians in the metazoan tree, and our results are consistent with this hypothesis.

39. The manuscript succeeds in the task of identifying novel SR in cnidarians and suggesting a role for them in the cnidarian-dinoflagellate symbiosis and I'd like to compliment the authors for that. Thank you

40. Although the manuscript might be ‘self-contained' in it's present form, it would be interesting to have a larger discussion in the effects between the cnidarian immune system, the production of NO and the disruption of symbiosis under oxidative stress. This is a great idea for a larger study or a review paper.

41. Would the oxidation burst in the dinoflagellate under heat stress causes the failure of the symbiont to regulate the host immune system? Possibly – this is related to the idea that symbionts are cloaked in the host when healthy and not detected – tolerated. But when they get stressed, they signal their presence and the host mounts an attack

42. It would be also interesting to evaluate the presence of the set of SR in non-cnidarian hosts. Riesgo et al (2014) presents transcriptomic resources for both symbiotic and aposymbiotic sponge Cliona varians. It would be interesting to include it in the bioinformatic analysis presented.

Riesgo, A., Peterson, K., Richardson, C., Heist, T., Strehlow, B., McCauley, M., Cotman, C., Hill, M. and Hill, A., 2014. Transcriptomic analysis of differential host gene expression upon uptake of symbionts: a case study with Symbiodinium and the major bioeroding sponge Cliona varians. BMC genomics, 15(1), p.1 Of course this would be very interesting to explore in the future!