

We thank Referee #1 and #2 for taking the time to review our discussion paper. We understand the paper in its present form, reveals some inconsistencies and confusing interpretations partially due to an apparent lack of “story-line” and incoherent flow between different sections and experiments reported. This will be reconsidered and carefully amended in a revised version of the manuscript but readers and public audience should be aware that the work presented is of descriptive nature and not aiming at the characterization of physiological reactions and molecular mechanisms in light of differences found in gene expression between Menez Gwen and Lucky Strike *B. azoricus* mussels. In the contrary, our manuscript was aiming at the finding of “signatures” or “markers” of descriptive nature, supported by differences found at gene expression levels, in-situ hybridization results and 16S amplicon sequencing results. All three have the potential to show differences of such markers or signatures pointing at the geographical origin of both Menez Gwen and Lucky Strike populations and interpret our results in light of our long-standing knowledge of immune genes in *B. azoricus* and more recently our microbial community studies from gill tissues.

We believe the paper was harshly misevaluated due to an error on our part in writing the probe sequences used in in-situ hybridization experiments. We have in most of the cases and several years, used oligonucleotide probes aimed at nuclear genes instead of ribosomal genes. While the so called “correct general “standard FISH probes have been extensively popularized, and Duperron’s probes are well known for targeting 16S rRNA, nothing really impedes someone of using other probes of different sequences and test them in FISH experiments as they might work just as well. Not only we wanted to use probes targeting nuclear genes (this somewhat novel but it has now become more and more in use, especially in Nicole Dubilier’s lab) we wanted also to base our probes on bacterial sequences that we have revealed from our own transcriptome studies in *B. azoricus* not someone else’s sequences. Ribo probes work as well as cDNA probes as long as they follow the rules of base complementarity, RNA integrity (whether or not the target RNA is intact) and are targeting coding sequences since this is what we wanted to target, after all, expressed mRNA. In the present study we wanted to target the MMO and sulfur oxidation genes and used probes that were designed to target the respective nuclear genes:

MMO- CACTAACTATGCTAACCGCGATGTCA

SOX- CGACTAGGAGCACATCTATTAGGTTT

The sequenced for our MMO probe design came from our sequence >mussel_c5320 length: 933 **methane monooxygenase protein A** [Methylococcaceae bacterium SF-BR] ensued from our transcriptomics studies already published and referred in the discussion paper as Bettencourt et al. 2010 BMC Genomics BMC Genomics, 11, 559, doi:10.1186/1471-2164-11-559, 2010 and Egas et al. 2012 Mar. Drugs, 10, 1765–1783, 2012.

The sequence for our SOX probe design came from our sequence >mussel_c3834 length: 922 **sulfur oxidation protein SoxY** [Sulfurovum sp. NBC37-1]

Ensued form the same published studies as above

A BLAST search in NCBI confirmed that our complementary sequence resulted in hits within *Bathymodiolus* MMO gene

Bathymodiolus brooksi gill symbiont clone GoM_Chap_pmoA_2.1
particulate methane monooxygenase A (pmoA) gene, partial cds

Sequence ID: [gb|JN021262.1](#) Length: 467 Number of Matches: 1
 Related Information
 Range 1: 58 to 83 [GenBankGraphics](#) Next Match Previous Match [First Match](#)
 Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand	Frame
52.0 bits(26)	7e-05()	26/26(100%)	0/26(0%)	Plus/Minus	

Features:
 Query 1 GTGATTGATACGATTGGCGCTACAGT 26
 |||
 Sbjct 83 GTGATTGATACGATTGGCGCTACAGT 58

[GenBankGraphics](#) Next Previous [Descriptions](#)
 Endosymbiont of Bathymodiolus puteoserpentis partial pmoA gene for particulate methane monooxygenase subunit A, sequence ID #7986
 Sequence ID: [emb|FR865039.1](#) Length: 471 Number of Matches: 1
 Related Information
 Range 1: 31 to 56 [GenBankGraphics](#) Next Match Previous Match [First Match](#)
 Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand	Frame
52.0 bits(26)	7e-05()	26/26(100%)	0/26(0%)	Plus/Minus	

Features:
 Query 1 GTGATTGATACGATTGGCGCTACAGT 26
 |||
 Sbjct 56 GTGATTGATACGATTGGCGCTACAGT 31

A BLAST search in NCBI confirmed that our sequence >mussel_c3834 length: 922 **sulfur oxidation protein SoxY** [Sulfurovum sp. NBC37-1] from which our SOX in-situ probe was designed is indeed matching proteins hits containing the SOX Y domain from SOX Y superfamily

[GenPeptGraphics](#) Next Previous [Descriptions](#)
 sulfur oxidation protein SoxY [Sulfurovum sp. AR]
 Sequence ID: [ref|ZP_10062574.1](#) Length: 155 Number of Matches: 1
 Related Information
 Range 1: 6 to 155 [GenPeptGraphics](#) Next Match Previous Match [First Match](#)
 Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
205 bits(522)	2e-62()	Compositional matrix adjust.	119/150(79%)	135/150(90%)	0/150(0%)	+1

Features:
 Query 301
 FLKSICAasavvatvpsllvaKDAPKGGNALS YDAAvvtitggkvvtgSDKIKLTVPEI 480
 F+KSICAASAV ATV+PS L AK+APKGGN LSYDAAV ITGGK V SDK+
 LTVPEI
 Sbjct 6
 FIKSICAASAVAATVTPSALFAKEAPKGGNVLSYDAAVAAITGGKAVADSDKVNLTVP EI 65

Query 481
 AENGAVVPVKVNVESPMTDADYVKAIHVLTTKNSNARCADVMLTPLNGKGYFATRVLGG 660
 AENGAVVPVKV+V+ PM + +YVKAIHVL+TKN
 NARCADVMLTPLNGKGYFATR+KLGG
 Sbjct 66
 AENGAVVPVKVDVDHPMEENNYVKAIHVLSTKNGNARCADVMLTPLNGKGYFATRIKGG 125

Query 661
 TQDVVALVEMSDGSFLRAAKPVKVTIGGCG 750
 TQDVVALVE+S+G+F+++AK VKVTIGGCG
 Sbjct 126
 TQDVVALVELSNGTFIKSAKSVKVTIGGCG 155

[GenPeptGraphics](#)[Next](#)[Previous](#)[Descriptions](#)

sulfur oxidation protein SoxY [Sulfurovum sp. NBC37-1]

Sequence ID: [ref|YP_001357815.1](#)|Length: 159Number of Matches: 1

Related Information

[Gene](#)-associated gene details

Range 1: 32 to 159[GenPeptGraphics](#)[Next Match](#)[Previous Match](#)[First Match](#)

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
178 bits(452)	7e- 52()	Compositional matrix adjust.	90/128(70%)	105/128(82%)	0/128(0%)	+1

Features:

Query 367

KDAPKGGNALSYDAAvvtitggkvvtgSDKIKLTVPEIAENGAVVPVKVNVESPMTDADY 546
K PKG NALS +AA+ ITGGK SDK+KLTVPEIAENGAVVPVKVNV+ PM +

+Y

Sbjct 32

KAVPKGPNALSVEAAIDAITGGKGAKESDKVKLTVPEIAENGAVVPVKVNVDPHMEEGNY 91

Query 547

VKAIHVLTTKNSNARCADVMLTPLNGKGYFATR.VKLGGTQDVVALVEMSDGSFLRAAKPV 726
VKAIHVL KN N+RC DVMLTP NGK YFATR+KLG TQ+V+ + E+SDG+F++AAK

V

Sbjct 92

VKAIHVLAAKNGNSRCVDVMLTPANGKAYFATRIKLGSTQEVIGVAELSDGTFIKAASV 151

Query 727 KVTIGGCG 750
KVTIGGCG

Sbjct 152 KVTIGGCG 159

So in conclusion, our probes are theoretically good for FISH experiments and results obtained do show signal specificity. The auto-fluorescence referred by the referee is simply not accurate. First of all it does not come from lipid droplets! But rather sugar granules that we have already figured out in 2008 and reported in a paper from that year Bettencourt et al. 2008 **Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology** May 2008, Volume 150, Issue 1, Pages 1-7. The ALEXA fluorochromes are the finest available in the market and their spectrum of absorbance and emission are well defined and would only being visualized with appropriate filters from the fluorescent microscope such as ours, a Leica DM6000. Auto-fluorescence is only visualized in the present case at UV emission well below the spectrum used by our ALEXA Fluorochromes...hence the specificity!

The following comment by the anonymous referee #1 raises some concerns

The cDNA library used for qPCR of bacterial genes was inappropriate as this was subjected to poly-A selection, a treatment designed to remove ribosomal RNA, but which also removes bacterial mRNA. For this reason, the V6 sequencing experiment also has little value, as the PCR products for sequencing were amplified from the same cDNA library

We have devoted 1 full-length article to this question apart from our work published in 2010 Bettencourt et al. BMC Genomics in which we described in detail how the cDNA library was built and how we ended up with 3000 bacterial genes sequences which

made through the poly-A selection during the course of library preparation and prior to 454 sequencing. Bacterial mRNA do not have poly-A tails, right, but it is not true that bacterial mRNA will be removed after poly-A selection, actually some bacterial RNAs are poly-adenylated too. The point here is not to debate whether or not our initial cDNA was “contaminated” with bacterial mRNA but to deal with the fact that in our initial work on the transcriptome sequencing of *B. azoricus* gill tissues (vide Bettencourt et al. 2010 BMC Genomics) some 3000 cDNA sequences were revealed pointing at functional bacterial genes that were subjected to the MG RAST, the Metagenomics RAST server, an automated analysis platform for metagenomes providing quantitative insights into microbial populations based on sequence data. This is a fact that cannot go unnoticed and that prompted us to dedicate another work published in Marine Drugs by Egas et al. 2012 with the title “The Transcriptome of *Bathymodiolus azoricus* Gill Reveals Expression of Genes from Endosymbionts and Free-Living Deep-Sea Bacteria”. I do agree with the anonymous reviewer in that the V6 sequencing experiment also has little value, given the possibility that cDNA libraries were poorly represented by bacterial mRNA, however it was not clearly written or stated in our paper that both LS and MG cDNA libraries were obtained by using random primers and not oligo-dT during the process of reverse-transcription and that would significantly change the outcome of our analyses. The same total RNA was used but reverse-transcribed differently, using random primers instead of oligo-dT. Moreover, the 16S amplicon sequencing was strictly meant for gill’s microbiome structure analyses, not for quantifying bacterial gene expression.

Furthermore, the cDNA library used for qPCR of bacterial genes might have been inappropriate to the reviewer’s view, as this was subjected to poly-A selection but it was primarily generated for host gene expression studies, nonetheless given our previous results on the transcriptome sequencing of *B. azoricus* gill tissues and the bacterial genes that were revealed then, we felt compelled to pursue these qPCR experiments even with the same cDNA libraries. It is not uncommon that mispriming events do occur with RNA species and often cases, even ribosomal RNA will be misprimed with oligo-dT and reverse-transcribed into cDNA. rRNA is expressed at a very high level, so even a little leaky priming by nonspecific priming would work rather well, but not as well as the random priming. Such abundant RNA species (rRNA) almost invite mispriming merely by being highly present and thus a small amount will always be reverse transcribed into cDNA, including bacterial mRNA, because of normal, expected low-level RT mis-priming events... One has to assume in the present case that bacterial mRNA was highly present in our gill total RNA extractions!

Finally I would like to address the immune genes that were referred by the anonymous reviewer as not being validated or shown to respond to infection or microbial stimulation. The reason why we did not present evidence of gene expression upon immune stimulation is because we have submitted another paper to the same BG special issue that is regarded as a companion paper and dealing precisely with this subject. It

has been submitted by Martins et al. from my group with the following title “Finding immune gene expression differences induced by marine bacterial pathogens in the Deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*”. However, I will be glad to include in the present paper, results from an infection experiment using *Vibrio parahaemolyticus* and *Flavobacterium* that was conducted in parallel, for validation purposes during the course of these studies (see below)

