Fabrication - probably unintentional

8 messages

Robert Karlsson <robertkarlss.se@gmail.com> Wed, Dec 21, 2022 at 14:40
To: research-integrity@glasgow.ac.uk

Dear Research Integrity team,

My name is Robert. I am biochemist from Sweden.

I have well-founded suspicions of research misconduct in University of Glasgow.

May I provide you with more informations?

Kind regards,

Robert

Research Integrity <research-integrity@glasgow.ac.uk> Wed, Dec 21, 2022 at 14:47
To: Robert Karlsson <robertkarlss.se@gmail.com>

Dear Robert,

Thank you for your email. Please can you send me the full details so that we can look into this further.

I would like to highlight that due to the time of year and staff already on annual leave we may not be able to properly look at the information until the University reopens after the holidays in January.

Best wishes

Amanda

From: Robert Karlsson <robertkarlss.se@gmail.com>
Sent: Wednesday, December 21, 2022 1:40 pm
To: Research Integrity <research-integrity@glasgow.ac.uk>
Subject: Fabrication - probably unintentional

[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Wed, Dec 21, 2022 at 15:19

To: Research Integrity <research-integrity@glasgow.ac.uk>

Dear Amanda,

OK. I understand.

I have reasonable suspicions that Professor Emma Thomson and Professor David Robertson use fabrication and manipulation inside University of Glasgow. They fabricate sars-cov-2 genome by experiment setup intentionally or unintentionally.

Please request the laboratory records, i.e. the proof of applied and documented negative control by Professor Emma Thomson and her team. The burden of proofs is on her team.

The negative control is very useful to exclude or prove fabrication by experiment setup in accordance with good laboratory practice and laws of logic.

The negative control: Use of RNA from uninfected cells/supernatant (cell cultures) treated in the same way as infected cell cultures but without any virus, and attempt of assembly of a virus genome with the reads derived from the RNA.

I am going to send the correspondence (PDF) with prof. Emma tomorrow, via this email.

Kind regards,

Robert
[Quoted text hidden]

Research Integrity <research-integrity@glasgow.ac.uk> Wed, Dec 21, 2022 at 15:32
To: Robert Karlsson <robertkarlss.se@gmail.com>

Dear Robert,

Many thanks for this initial information. I will read through all of the PDF information that you send tomorrow. As previously stated I am likely unable to take the information of the allegation forward until into the new year but I will keep you updated as and when is necessary.

Best wishes,

Amanda

[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Thu, Dec 22, 2022 at 09:40
To: Research Integrity <research-integrity@glasgow.ac.uk>

Dear Amanda,

On the basis of the scientific publications of the mentioned scientists as well as the correspondences, an indication of the absence of applied and documented control within the laboratories inside the University of Glasgow is generated. Also, the scientists could use and ask for the results of the control carried out by a (any appropriate) colleague or a team.

Unfortunately, scientific journals do not initiate investigations into potential cases of fabrication within the University of Glasgow.

The negative control is no matter of discussion but documentation, basic logic and good laboratory practice. If you would try to describe the control without using the word "negative control" at all, you would get the exact name of the control by scientist (these scientists are from Belgium and Netherlands). I have 2 examples for this. I could share the examples if neccessary. This means that scientists are aware of the control. An even worse case would be the absence of awareness.

One example of logical and neccessary negative control:

When a sterile swab comes into contact with the human mucosa, the consequence of the contact is the introduction of epithelial cells on the swab. This means that the negative control should contain RNA extracted from human epithelial cells. VTM (virus transport medium with antibiotics and antifungals) should be used in the generation of the negative control also. Considering that human epithelial cells are contaminated with human microbiome, it is necessary to use RNA from epithelial cell cultures. Epithelial cell cultures are virus-free. So, this control is very useful to exclude or prove generation of sars-cov-2 genome without the presence of sars-cov-2 genome.

The use of the wuhan-hu-1 reference genome (MN908947) is particularly worrying because it is based on manipulation. I don't know if Professor Emma Thomson and her team use manipulation-based genome intentionally or unintentionally.

The manipulation is explained in the attachment. The correspondences are in attachment.

I hope that in the New Year you/your team will find or initiate a suitable solution or an idea for solution even though this case looks a little scary.

Best regards and Happy Holidays,

Robert
[Quoted text hidden]

Research Integrity <research-integrity@glasgow.ac.uk> Thu, Dec 22, 2022 at 11:09
To: Robert Karlsson <robertkarlss.se@gmail.com>

Dear Robert,

Many thanks for this information and attached PDF's. Can you please provide me with any emails that you received from David Robertson as these seemed to be missing from what you sent me. I only seem to have your responses.

Many thanks,

[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Thu, Dec 22, 2022 at 11:22
To: Research Integrity <research-integrity@glasgow.ac.uk>

Dear Amanda,

He was invited by prof. Emma to explain the manipulation (or mistake) in the field of bioinformatics. They work together (shared responsibility). Unfortunately, he did not answer. Maybe he's scared or in panic. If he is out of the office for a long time, there is an option of an automatic reply with the exact number of days of absence from the office.

Welcome,

Robert
[Quoted text hidden]

Research Integrity <research-integrity@glasgow.ac.uk> Thu, Dec 22, 2022 at 11:26
To: Robert Karlsson <robertkarlss.se@gmail.com>

Dear Robert,

Many thanks for the clarification.

[Quoted text hidden]

Research integrity concerns at the University of Glasgow

6 messages

Andrew Roe <Andrew.Roe@glasgow.ac.uk> Thu, Feb 2, 2023 at 20:12
To: robertkarlss.se@gmail.com <robertkarlss.se@gmail.com>
Cc: Research Integrity <research-integrity@glasgow.ac.uk>

Dear Robert

Please let me introduce myself. I am Professor Andrew Roe, and act as the Research Integrity Champion for the College of Life Sciences. The research integrity team has passed your emails and concerns for review. I am keen to resolve this to everyones satisfaction and I can assure you we take these allegations very seriously.

I have read the correspondence and attachments that you provided and had a few questions. I **apologise if some of these appear obvious to you, but it is very important that I fully understand the allegations so that we can closely assess the evidence.** For clarity I have written some very key questions in blue font below to make our discussion absolutely clear. I hope that helps.

In your previous emails you say

"They fabricate sars-cov-2 genome by experiment setup intentionally or unintentionally."

Can you please explain to me in simple terms the basis to why you believe the SARS-COV-2 has been manipulated? From the emails, it seems that you tried to replicate the assembly but got a different or conflicting result?

Did you receive any correspondence from the bioinformatics team in Glasgow who were involved in the assembly?

Secondly, you emphasise the important negative controls. Of course all experiments should be carefully controlled but what is it about this case that makes you believe they have performed the wrong controls? The email trail seems to indicate that a series of negative controls were run but you feel these were incorrect or inadequate? Given the paper was accepted in nature

Microbiology, then it has been independently peer reviewed by relevant experts. Why would they not raise this concern if it is very clear?

Finally, again perhaps a silly question from me (I am a bacteriologist!)- surely the genome of this virus has now been sequenced many more times? This is simple enough to perform so surely there is independent evidence of the genome sequence that can be directly compared?

Many thanks for your patience whilst we investigate this case

Kindly

Andrew

Robert Karlsson <robertkarlss.se@gmail.com> Sat, Feb 4, 2023 at 12:36
To: Andrew Roe <Andrew.Roe@glasgow.ac.uk>

Dear Andrew,

I am happy to provide more informations:

1. The raw reads of MN908947 are manipulated. Some reads are removed from the original file and some reads are inserted in the file. So, raw reads are not raw but manipulated and the assembly is generated with manipulated data. So, the assembly is manipulated. Generation of MN908947 is described in this study: "A new coronavirus associated with human respiratory disease in China". The paper was accepted in Nature which indicates that the submission went through peer review. MN908947 is generated with default parameter settings (Megahit v1.1.3). However, it is not possible to reproduce the result (assembly = the longest contig = 30,474 nt) with default parameter settings and the published raw reads. I tried to reproduce the result (assembly) and it is not possible. So, the assembly is manipulated. There are more than 300,000 SARS-CoV-2 studies and no one tried to check the validity of MN908947. Thousands and thousands of scientists used this construct (de novo assembly) but no one checked this construct in the sense of the existence of manipulation or fabrication or mistakes. All of the scientists just believe that MN908947 is legitimate and use it.

2. I didn't receive any correspondence from David Robertson (bioinformatics).

3. If you put RNA from a human into a sucrose solution, you can say that the RNA comes from a human. However, if you mix RNA from humans with RNA from cell cultures and from

bovine sources (study: "Extracellular small non-coding RNA contaminants in fetal bovine serum and serum-free media"), a problem arises: You cannot determine the origin of the RNA . Bioinformaticians use reference genomes to map the reads, but these reference genomes are also created from mixed RNA samples without logical and necessary negative control.

The negative control I described is based on basic logic.

You can seek responsibility from prof. Emma and her team, peer review, as well as accreditation bodies. For example, you can check whether the methods are validated with the described negative control. I personally don't care who did and documented the control in detail. I would be satisfied if detailed documentation exist. If it doesn't exist, someone should URGENTLY conduct and document the control. This control must be disclosed. I am very worried about this.

Also, if prof. Emma never did and documented the control, maybe there is suitable documentation from any appropriate virologist or bioinformatician in the world? She could ask her colleagues.

Those would be some of the options for disclosure of the negative control.

The most commonly used negative control for sequencing is nuclease-free water known as "blank". This type of control excludes the possibility of contamination. However, this control does not exclude the possibility of fabrication by experiment setup. There is one enormous gap in the controls used by prof. Emma - the described negative control has never been conducted and documented. Maybe I am wrong?

The neccessary negative control: Use of RNA from uninfected cells/supernatant (cell cultures) treated in exactly the same way as infected cell cultures but without any suspected virus, and attempted assembly of the virus genome with the reads derived from the RNA.

4. The SARS-CoV-2 genome has been sequenced many, many times around the world, most often with an amplicon-based WGS approach, but without the necessary negative control. The result of the negative control would be raw reads available and percent of alignment to SARS-CoV-2 genome. The number of PCR cycles used in the control should be documented.

I hope this helps.

Best regards,

Robert
[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Sat, Feb 4, 2023 at 20:13
To: Andrew Roe <Andrew.Roe@glasgow.ac.uk>

Dear Andrew,

This table is useful as an example of documenting controls. You can see how much RNA was extracted from the supernatant and how much from the cells. These are human epithelial cells without the presence of any virus. These are unpublished results.

You can ask prof. Emma for appropriate documentation. There could be one more table with the exact number of sars-cov-2 reads from the controls.

The table is available in attachment, pdf file.

Best,

Robert
[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Sun, Feb 5, 2023 at 09:22
To: Andrew Roe <Andrew.Roe@glasgow.ac.uk>

Dear Andrew,

This explanation could be useful:

When using cell cultures in virology, the negative control for CPE (cytopathic effect) implies the treatment of cell cultures in the same way as for infected cultures, but without any virus (patient sample).

According to the latest knowledge, the result of experiments in virology is a complete construct (whole-genome assembly) generated after virus isolation in cell cultures and sequencing of the supernatant, or generated directly from the patient sample without any isolation.

So, according to the latest knowledge, when using cell cultures in virology and sequencing and assembly of a RNA virus, the neccessary and logical negative control implies:

Use of RNA from uninfected cells/supernatant (cell cultures) treated in the same way as infected cell cultures but without any virus, and attempt of assembly of a virus genome with the reads derived from the RNA.

Best,

Robert
[Quoted text hidden]

Andrew Roe <Andrew.Roe@glasgow.ac.uk> Sun, Feb 5, 2023 at 12:39
To: Robert Karlsson <robertkarlss.se@gmail.com>
Cc: Research Integrity <research-integrity@glasgow.ac.uk>

Dear Robert

Thanks for the additional information and for helping clarify those areas.

Can I ask one further question please.

You refer to the manipulation of the MN908947 coronavirus genome and say some reads are removed. This should be easy to check and I can do this myself. However, this study was published by a Chinese group and has nothing to do with University of Glasgow. Why is this manipulation relevant to the group in Glasgow?

1. The raw reads of MN908947 are manipulated. Some reads are removed from the original file and some reads are inserted in the file. So, raw reads are not raw but manipulated and the assembly is generated with manipulated data. So, the assembly is manipulated. Generation of MN908947 is described in this study: "A new coronavirus associated with human respiratory disease in China". The paper was accepted in Nature which indicates that the submission went through peer review. MN908947 is generated with default parameter settings (Megahit v1.1.3). However, it is not possible to reproduce the result (assembly = the longest contig = 30,474 nt) with default parameter settings and the published raw reads. I tried to reproduce the result (assembly) and it is not possible. So, the assembly is manipulated.

Thanks

Sent from my iPad
[Quoted text hidden]

Robert Karlsson <robertkarlss.se@gmail.com> Sun, Feb 5, 2023 at 12:57
To: Andrew Roe <Andrew.Roe@glasgow.ac.uk>

Dear Andrew,

Thank you for your question.

You should pay attention on Sat, Dec 3, 2022 at 14:33; the correspondence with prof. Emma:

"Your study says: "Illumina paired-end reads were aligned to the Wuhan-Hu-1 reference genome (MN908947.3) using bwa74""

So, the team of prof Emma is involved in the manipulation by alignment to MN908947. Intentionally or unintentionally, I dont know. I think they are just not aware, but I am very worried why her team didn't check the assembly used for alignment. And I am worried why there is no any explanation from her team.

Unfortunately, prof. Zhang from China didn't provide any response. So I am in position to find other ways. This is one of the solutions and I am grateful for your help, and maybe David Robertson could help also with your intervention. Something like that, for the problem in the field of bioinformatics.

Best regards

Robert
[Quoted text hidden]

*Confidential - Conclusion to Integrity Complaint

1 message

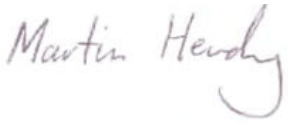
Research Integrity <research-integrity@glasgow.ac.uk> Thu, Mar 9, 2023 at 14:11
To: Robert Karlsson <robertkarlss.se@gmail.com>

Dear Robert,

Thank you for sharing your concerns regarding the virology researched by Prof Emma Thompson at the University of Glasgow. After further investigation, it appears that the major concern lies with the original Wuhan reference genome MN908947, first published in Nature. As it stands, there is no documented evidence that the assembly in question was wrong (Author Correction, editorial or another paper). However, as you believe it is incorrect, I would therefore encourage you to show the Nature editor where the assembly has been manipulated as they may then do their own investigation.

After speaking to several colleagues working in bioinformatics, we do not expect our scientists to check if an assembly is correct, as none of our immunologists would ever re-assemble the human genome. It is simply being used as a scaffold for further re-sequencing experiments. Therefore we do not believe any of our researchers have been involved in research misconduct and we will not be investigating any further.

Best wishes,



Professor Martin Hendry

Clerk of Senate and Vice Principal

Named Person for Misconduct Allegations

Email: research-integrity@glasgow.ac.uk