The biter being bit before the backlash. Bik keeps seeing copy and paste everywhere, so she finds some, but they are not !

Analyses on Elisabeth Bik's fraud allegations

Xavier Azalbert et Charles V – *France-Soir* 6 February 2024

In the wake of Elizabeth Bik's allegations of fraud, many of which are directed at the research center IHU-Méditerranée, we took an interest in her work. And as a matter of fact, one of her articles from 2022 contains many grey areas, and she is far from above suspicion.

For example, on the subject of her past experiences, she declares that she left uBiome in 2018. uBiome is an American start-up subject in 2019 to a resounding bankruptcy and an FBI investigation, forcing the two founders to leave the USA in a hurry and take refuge in Germany! Elizabeth Bik was the company's scientific director and wrote numerous articles, published with the uBiome founders. She did not ever request a retraction of these papers, even though she knew that the microbiota tests sold by the company didn't work. An investigation into the matter revealed substantial insurance fraud, as old samples were re-analyzed and billed to the insurance company a second time! As a conscientious scientist, Elisabeth Bik couldn't have **not been unaware of this**. The obscurity surrounding this situation has been confirmed by microbiome specialists. For them, "*Bik made a deal with the founders or federal investigators*".

Another grey area concerns **her real date departure date from uBiome**. Bik stated on X that it was in 2018, however, on January 9, 2019, she was still communicating about her participation in a health conference organized by the investment bank JP Morgan. Despite *France-Soir* reaching out to her on two occasions, she was not available to answer our questions...

A verification of Elisabeth Bik's fraud claims would do no harm. After all Bik claims she specializes in identifying photo fraud and photo duplicates. Therefore her analyses should be replicable. Using Charles V., expertise in computer graphics, we found **numerous inconsistencies in Bik's analyses**. Could this form part of reasons why she was unavailable to respond to our queries?

Hereafter, we **present the analyses conducted and the reasons why Bik is so wrong in her conclusions**, which could be linked solely to problems of definition of the images published. These errors yield several consequences as **they contradict Bik's fraud accusations** that she

made regarding many scientists, contributing to **casting doubt on their work and consequently muddy their reputations**.

Introduction to the verification analyses of BIK's claims

Elisabeth Bik specializes in "detecting studies that she believes are tainted by fraud". To this end, she uses the Pubpeer platform, which presents a series of tools designed to strengthen the traditional peer review process, notably by allowing comments from scientists and members of the public. A kind of social network for exchanging information on science.

On her website, Bik claims to be a "scientific integrity consultant" with "1069 retractions, 149 expressions of reservations, and 1008 Corrections (as of November 2023)". Bik specifically targets IHU-Méditerranée publications.

Therefore Charles V. looked into Bik's allegations concerning several IHU-Marseille articles. Let's look at the reanalysis of Bik's claimed fraud discovery.

1. **Erroneous conclusions on the first reanalyzed article**

Let's take a look at the article *"UUsing rpoB genetic analysis to detect and identify Bartonella species"* published on February 1st, 2001. Bik commented on Pubpeer on the **similarity in some areas** in a photo, implying that **this might constitute fraud**.

BIK'S INTEREST IN THE PAPER - *Using rpoB genetic analysis to detect and identify Bartonella species*

#1 Elisabeth M Bik commented March 2021

Figure 4C *Areas in the same photo look more similar than expected
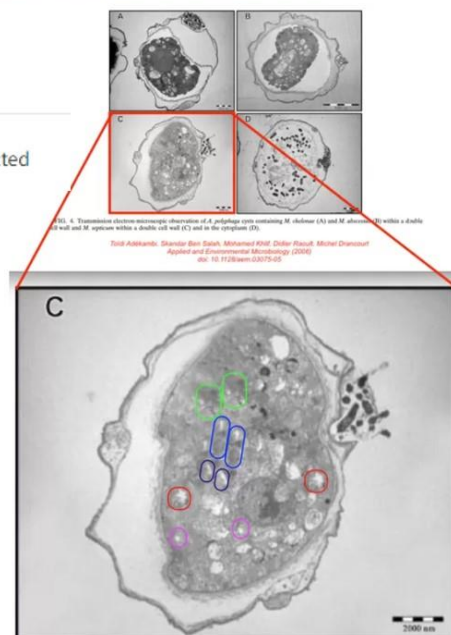
Ms. BIK's comments are as follows: "*In images 3 and 4:*

- Boxes of the same color show areas (some including stripes) that are more similar than expected.
- The green boxes show clear background transitions.
- The area above the stripes in lanes 1 to 5 of figure 4 also appears to show repetitiveness.*"

She adds that the "*image has been made darker to bring out more features. This is probably far from the case, but would the authors still have the original photos on gel?*"
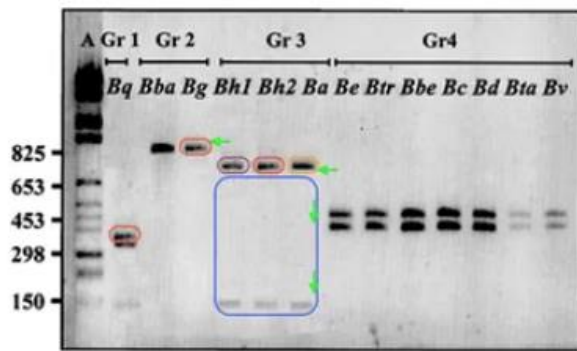
allowed easy discrimination of these strains. Thus, *ApoI* diges-
tion led to four different patterns. *B. quintana* was the only

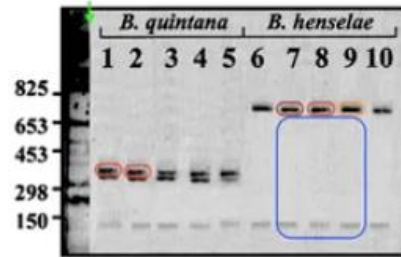FIG. 3. Restriction profiles obtained after digestion of a portion of
the *rpoB* gene with *ApoI*. Ethidium bromide-stained agarose gels of
ApoI restriction endonuclease digests of DNA amplified by using prim-
ers 1400F and 2300R are shown. Lane A, molecular mass markers
(marker IV; Boehringer). *Ba, B. alsatica; Bba, B. bacilliformis; Bbe, B.
berkhoffii; Bc, B; clarridgeiae; Bd, B. doshiae; Be, B. elizabethae; Bg, B.
grahamii; Bh1, B. henselae* Houston; *Bh2, B. henselae* Marseille; *Bq, B.
quintana; Bta, B. taylorii; Btr, B. tribocorum; Bv, B. vinsonii.* Numbers
on the left are in base pairs.

FIG. 4. *ApoI* digestion profiles of *rpoB* amplicons from either *B.
henselae* or *B. quintana*. Ethidium bromide-stained agarose gels of
ApoI restriction endonuclease digests of DNA amplified by using prim-
ers 1400F and 2300R are shown. Lane A, molecular mass (marker IV;
Boehringer); lanes 1 to 5, DNA extracts from blood of patients in-
fected with *B. quintana*; lanes 6 to 10, DNA extracts from lymph node
or pus aspirate samples from patients suspected of having cat scratch
disease and identified as *B. henselae*-positive samples. Numbers on the
left are in base pairs.

Verification of the first article challenged by Bik

It's time to check Bik's assertions. After all, in science, the important thing is replicability, but it's also a question of tools and methods. The verification process is explained below.

In order to verify the photo, we extracted the image challenged by Bik from the original PDF file to guarantee its authenticity. Its properties and characteristics are as follows: img60.png PNG 1960x1198 1960x1198+0+0 8-bit Gray 256c 471046B 0.000u 0:00.000

Bik claims, **without any real evidence, that two areas of the photo are duplicated**.

BIK CLAIMS THAT TWO ZONES COULD BE DUPLICATES

Bik claims that these 2 zones could be duplicated
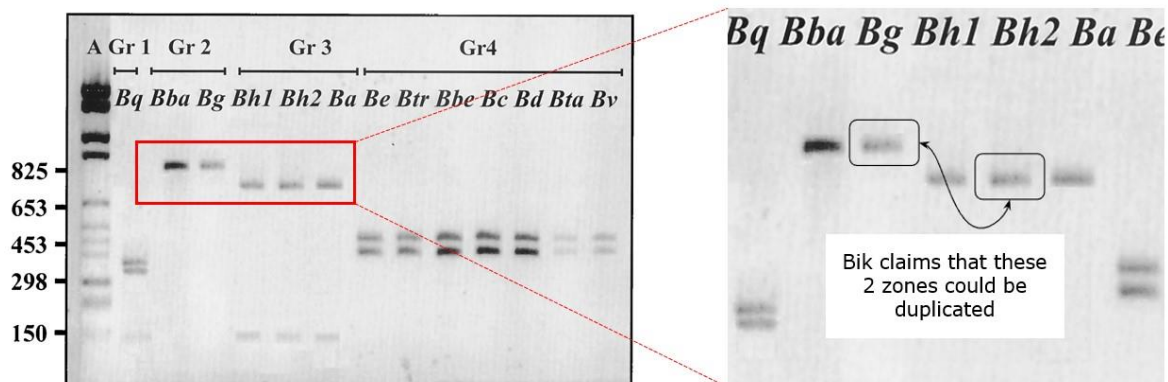
Using simple photo analysis tools, and after detailed examination, **Bik's argument is wrong**. This cannot be a duplication, but rather two areas which, although look visually similar, cannot be the result of a duplication.

The biter being bit. Let's take a look at Bik's **completely erroneous conclusion**, based on two methods.

A - First method: let's check whether "areas identified as duplicated" are really duplicated !

The photo on which there are allegedly duplicated areas is in the picture below.

Using image editing software (GIMP):

- The **Bg area** is extracted. This area is 84 pixels wide by 51 pixels high;
- A copy-paste of this area is duplicated on the photo to create a test area. This makes it possible to duplicate with certainty the area that is supposed to have been duplicated, the famous hypothesis put forward by Bik.

EXTRACTION OF THE PRESUMED DUPLICATED ZONE Bg
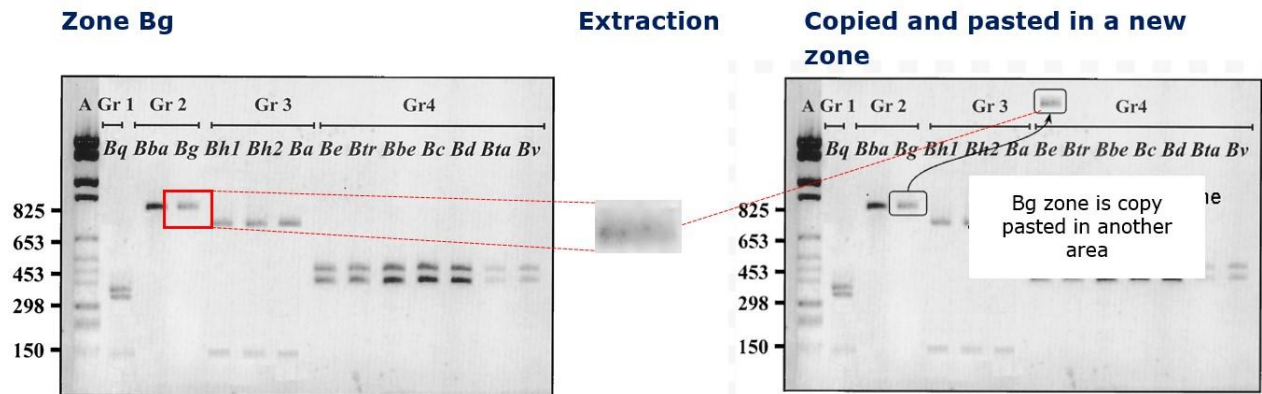
A Computer graphics software can be used to find out if there is a sequence (or pattern) in this photo. Given that part of the image has been duplicated on the picture, the software should technically find this sequence.

The recognition algorithm uses three parameters: the source image, the pattern image (the small excerpt of the area of interest) and a precision threshold ranging from 0 to 1 (1 being the exact image, 0 corresponds to matching everything in the entire photo).

At a 0.999 threshold, **both the pattern and the copy-pasted control are recognized**, and no other areas are detected. By lowering the threshold to 0.97, three zones are identified as identical. This is what Elisabeth Bik was able to detect and draw conclusions about duplication and therefore fraud.

DUPLICATE INDENTIFICATION WITH DIFFERENT PRECISION
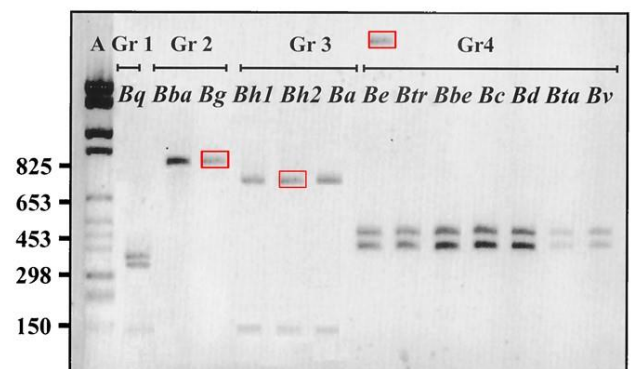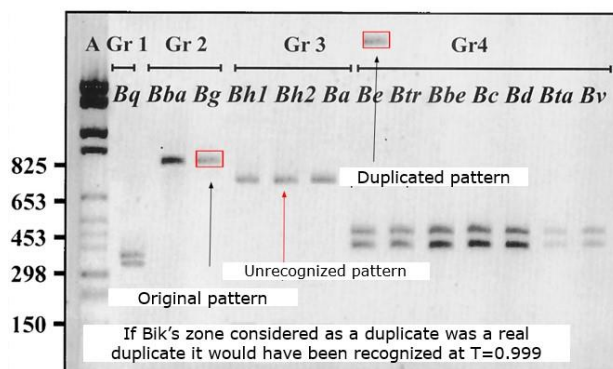
It would appear that she didn't go any further, which is rather **imprecise**, to say the least! There is a 3% difference between the source image and the new area found, which can be calculated from the differences between the thresholds used to identify the pattern (0.999 - 0.97) x 100 = 3%. In terms of pixels, i.e. the number of points used to represent the image, 3% of the pattern's area, i.e. 129 pixels, are different, or 3% of the difference between all the
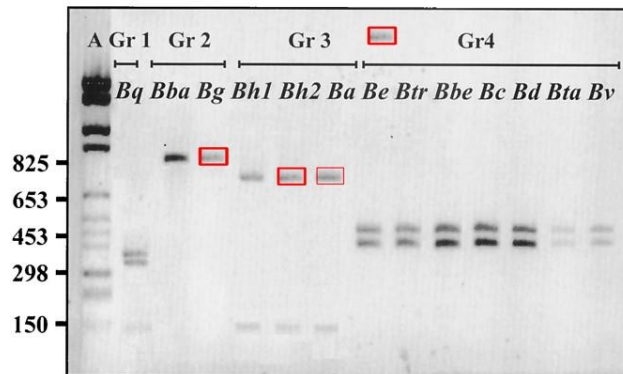
pixels in the image detected: this 3% variance **demonstrates the impossibility of a copy-paste operation**.

The analysis below shows that, at a 0.95 threshold, a fourth zone could have identical characteristics. At a 0.90 threshold , a multiplicity of zones appear.
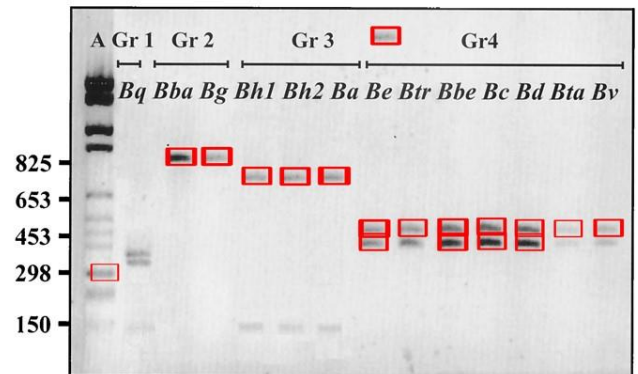
DUPLICATE ANALYSES WITH T = 0.95 and T = 0.90   France-Soir

T=0.95 – 4 zones are considered as duplicates   **T=0.90 – Many zones are considered as duplicates**

Discussion and analyses :

- At the .999 threshold, **only two zones are identical**. If the zones identified were really duplicated, as Bik claims, **we should have found three identical zones**. This is not the case. Bik therefore makes a **fundamental error** in not performing this analysis, which would have enabled her to verify that the zone she identifies as duplicated is not in fact so. She did not, as she should have done in quasi-exact science, try to reject her hypothesis.

- One problem could be the compression of the JPEG image in the PDF file. The verification was carried out by compressing the image file to 80% and the same result was obtained, which means that JPEG image compression plays no role in this situation.

- To take the analysis a step further, a 0.97 threshold search for "duplicated areas" was carried out in an image where the degree of certainty of the absence of copy-paste is close to 100%. Let's take a Nasa image, of Jupiter to be exact. Scanning the entire image, for each point a small rectangular area is taken and a clone search is launched, displaying in a red rectangle whether a clone is found at 0.97. With Jupiter, here's the result! So if BIK was telling the truth, she should also denounce the James Webb telescope images as fraud.

**ANALYSIS OF JUPITER – SEARCH FOR IDENTICAL ZONES –
MANY ARE IDENTIFIED**

Moreover, the colometry in Pubpeer is different from the original image in the PDF file, and Bik writes in her article that "*the images were further examined for evidence of duplication or image manipulation using the color adjustment tool in the Preview software on an Apple iMac computer*".

Ironically, if Bik turns off her screen, or sets the brightness to zero, she'll find a big copy and paste of black. Irony aside, a screen can be calibrated, and there are devices that allow you to do so.

In conclusion, these non-automated, non-reproducible image analysis techniques are **not scientific and are purely subjective.**

B - Methodology 2: Identifying highlights

Bik may not be using an automated, reproducible scientific approach to identify the frauds, but we felt it's important to test another method.

The second method used is that of **comparing remarkable points** in the image using a high-performance detection tool, FAST Feature Detector.

This tool is often used for point tracking in image stabilization. The assumption is: if two areas are similar, they have the same remarkable points.

(https://docs.opencv.org/4.x/df/d0c/tutorial_py_fast.html)

Here's the result obtained on the same image. The **result is incontrovertible**, confirming the hypothesis that there **has been no copy-paste as suggested** by the fraud hunter: between the original zone and the copied zone, the remarkable points are similar, whereas for the zone "imagined" by Bik, they are different not only in their cardinalities, but also in their positions. This is compatible with the first methodology

SECOND METHOD : REMARKABLE POINT ANALYSIS

France-Soir

Remarkable points (RP) identification

RP are identical : normal as true duplicate

Points are differing

No doubt, RP are differents through a cv2.FastFEatureDetector_create
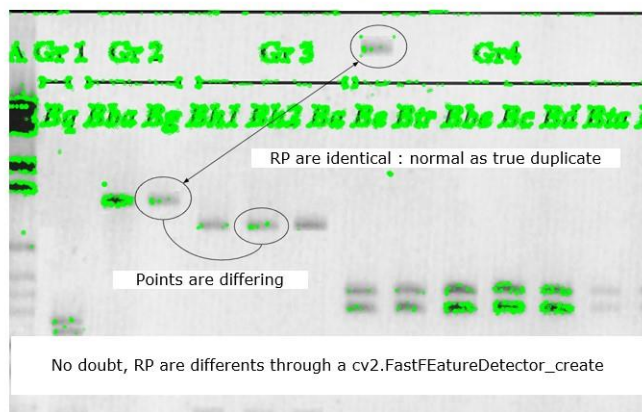
Visible differences - on the 3 zones original, real duplicate, zone considered by Bik as duplicated

Original | Real Duplicate | zone considered as duplicate
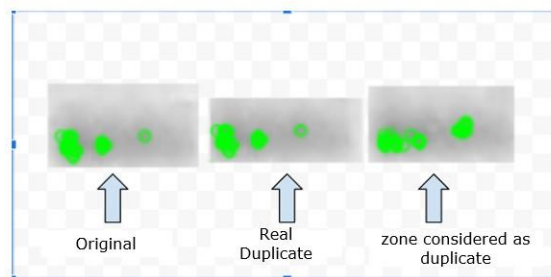
2. **Second verified article: Bik's conclusions wrong again**

For the second article identified as potentially having duplicated areas in an image (Survival of Environmental Mycobacteria in Acanthamoeba polyphagia), a simple analysis was carried out on the image, which may seem ridiculous in both methodological and scientific terms. The image was saved and opened in GIMP image editing software.
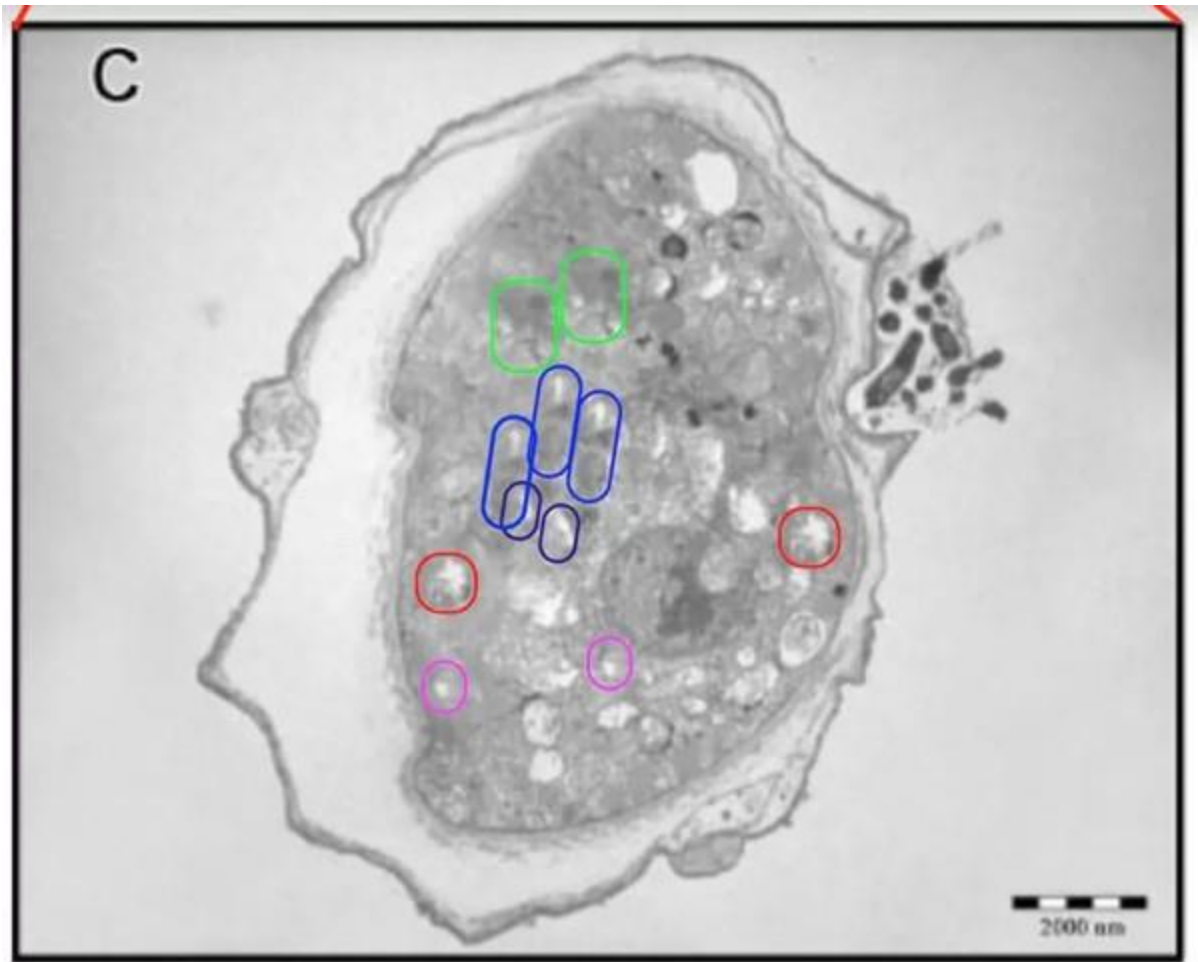
The first step is to zoom 1100% into the image - a naked eye can see that the areas are different.
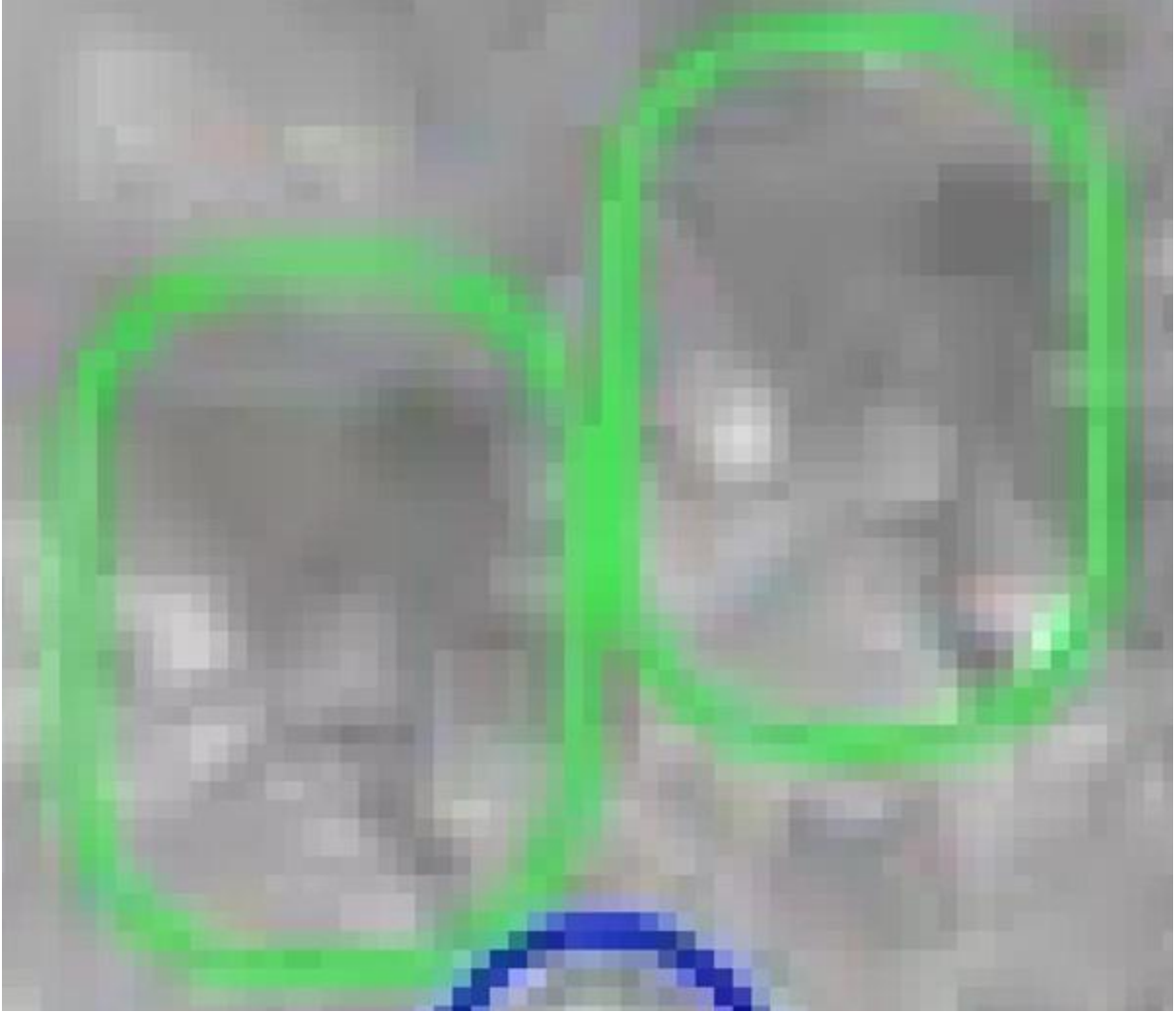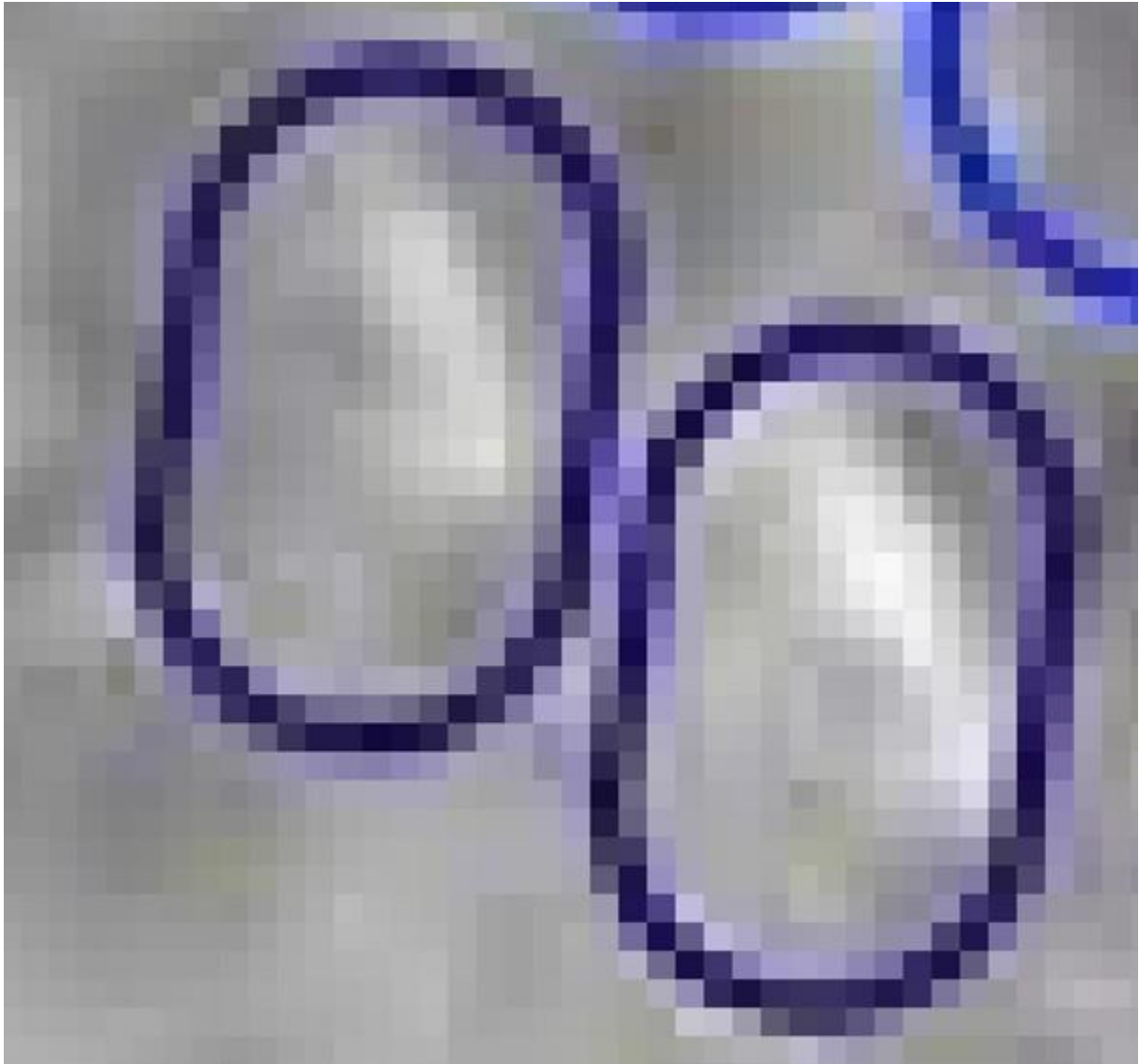
Zooming at 1600%. Same result.

At 2300% on the other zone, one can still see with the naked eye that the pixels are different.

3. **Third publication verification on an 2005 article** (Bartonella vinsoniiarupensis as an agent of blood-culture-negative endocarditis in humans). Bik also claims duplicated zones problems, which she reports on Pubpeer:

Figure 4 shows five panels that each represent different blots, incubated with different monoclonals.

- Red boxes: Lanes 1 from the MA4-C11 and MA8-F6 blots look remarkably similar. A potential splice can be observed in the MA8-F6 blot between lanes 1 and 5.
- Blue boxes: Lanes 1 from the MA2-E3 and MA2-A1 blots look remarkably similar. A potential splice can be observed in the NA2-A1 blot between lanes 1 and 3.
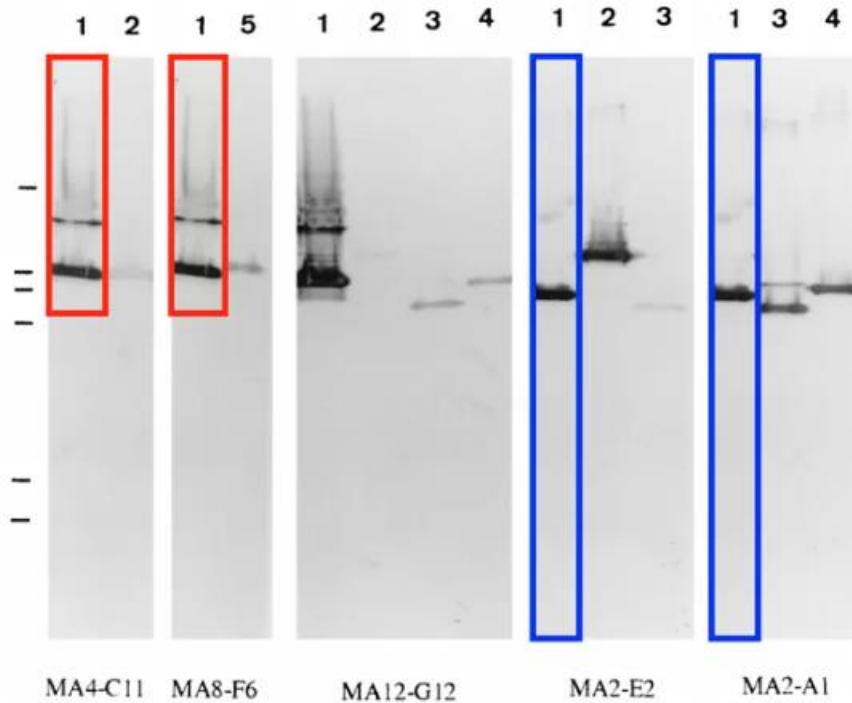
We verify Bik's analysis by demonstrating that the online service she used is not 100% reliable.
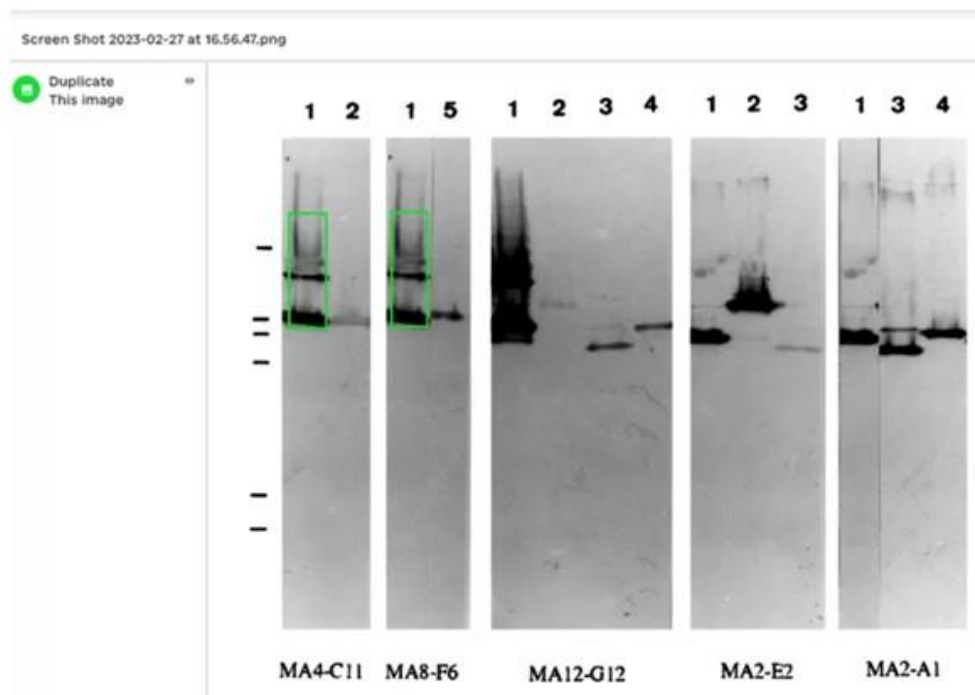
We create a duplicate of the zone that Bik considers identical by checking with the matching application used in the very first check: at the 0.999 threshold, **there is no duplicate zone identified**, but there is one at the 0.97 threshold. Déjà vu: same result, same conclusion.

DUPLICATES SEARCHES
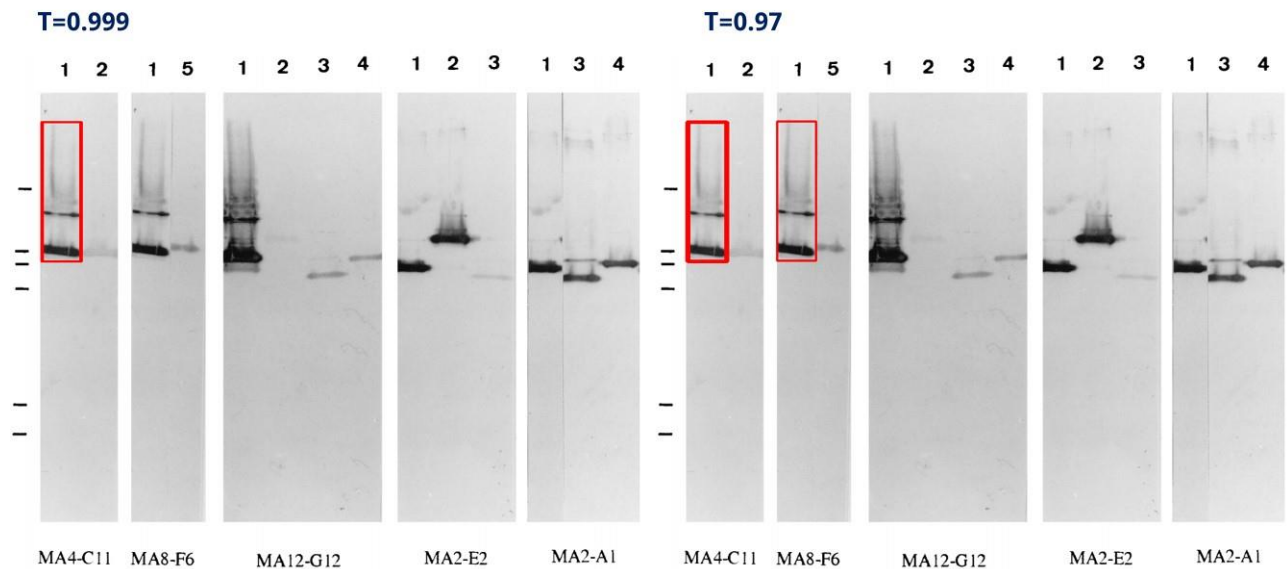
In response to her Pubperr comment and argument: "... Forensically also found the other duplication, as well as several false positives". However, using the FotoForensics software, with the default settings, **we don't get the same results** (illustration below). In science we should obtain results replication.

NON REPLICABILITY OF BIK'S ANALYSIS WITH PHOTO-FORENSICS

Image from pubpeer – Bik does not mention the software used
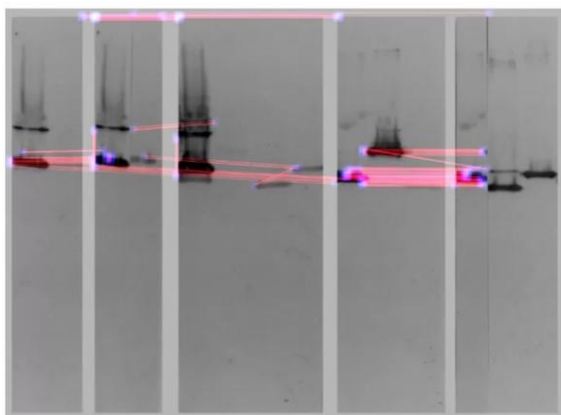
Image analyzed with photo-forencics
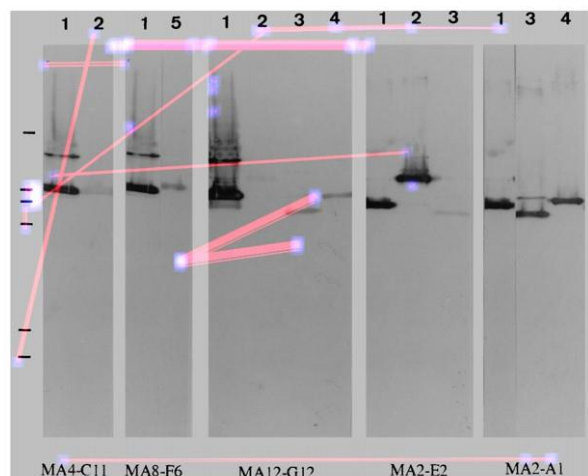
This is an unscientific analysis, as Bik doesn't indicate how she managed to achieve this result, which is therefore not replicable.

Sherlog can be used to display highlights: although geographically "fairly" close, they are different, and so are their coordinates.
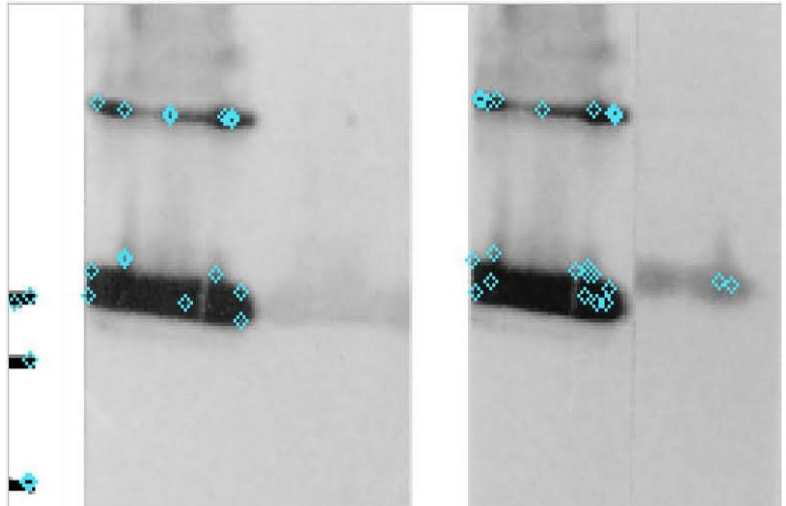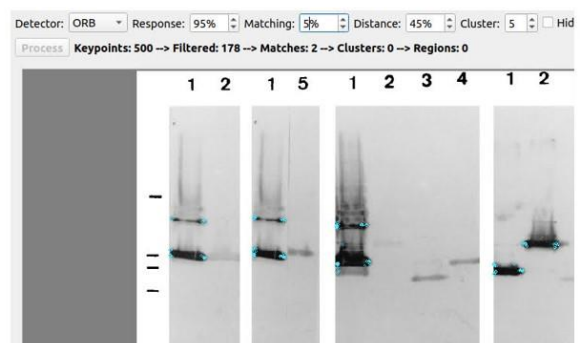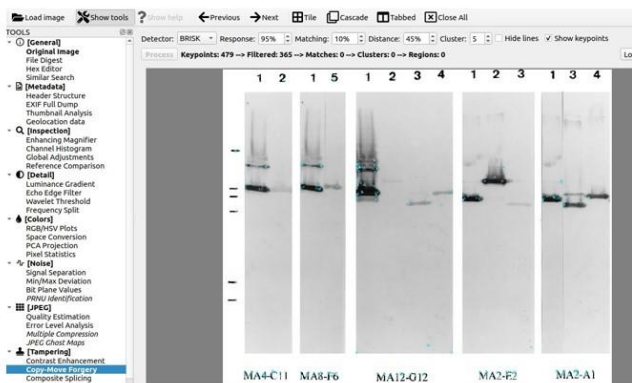

Sherlog offers several methods for detecting cut-and-pastes

With BRISK (Binary Robust Invariant Scale Keypoint): no detection. With the ORB Oriented FAST and Rotated Brief , with a sensitivity threshold of 5% matching, no copy found.
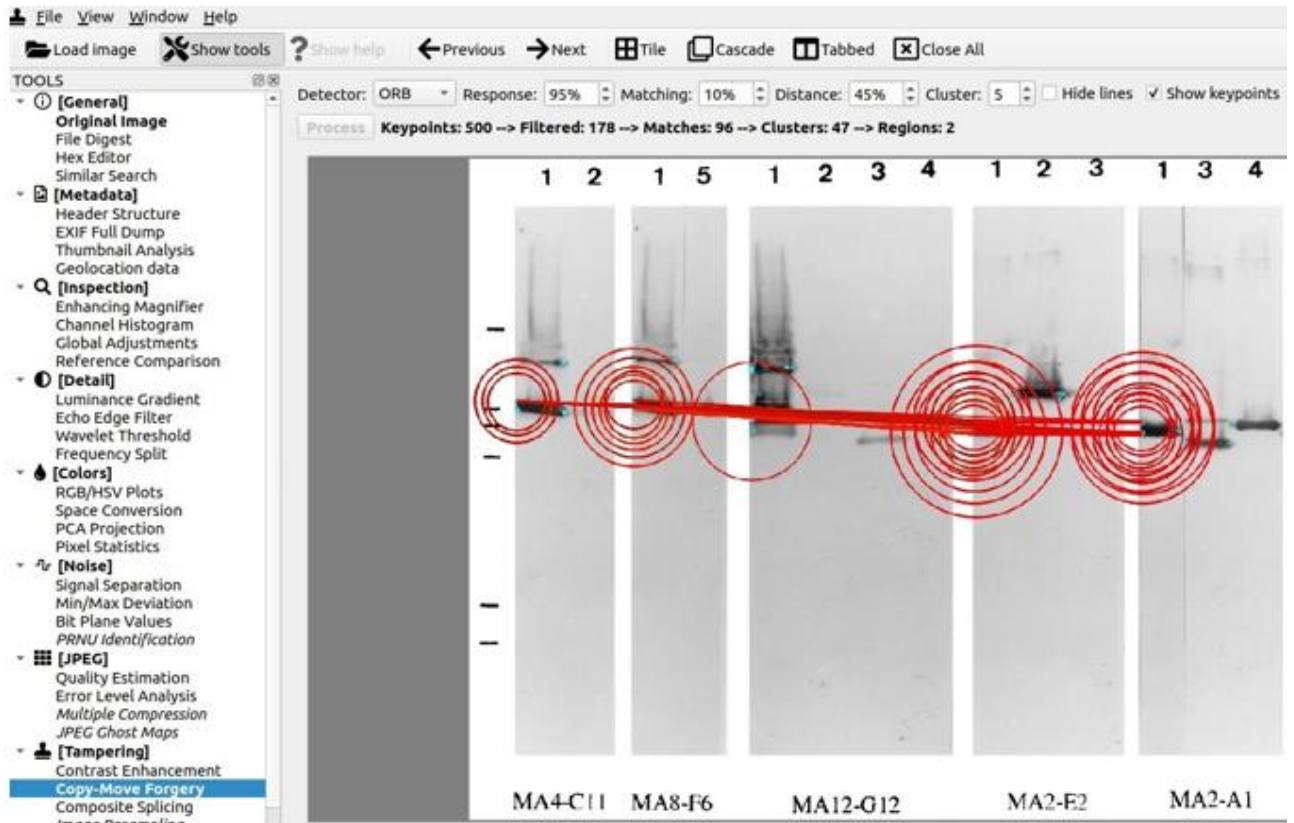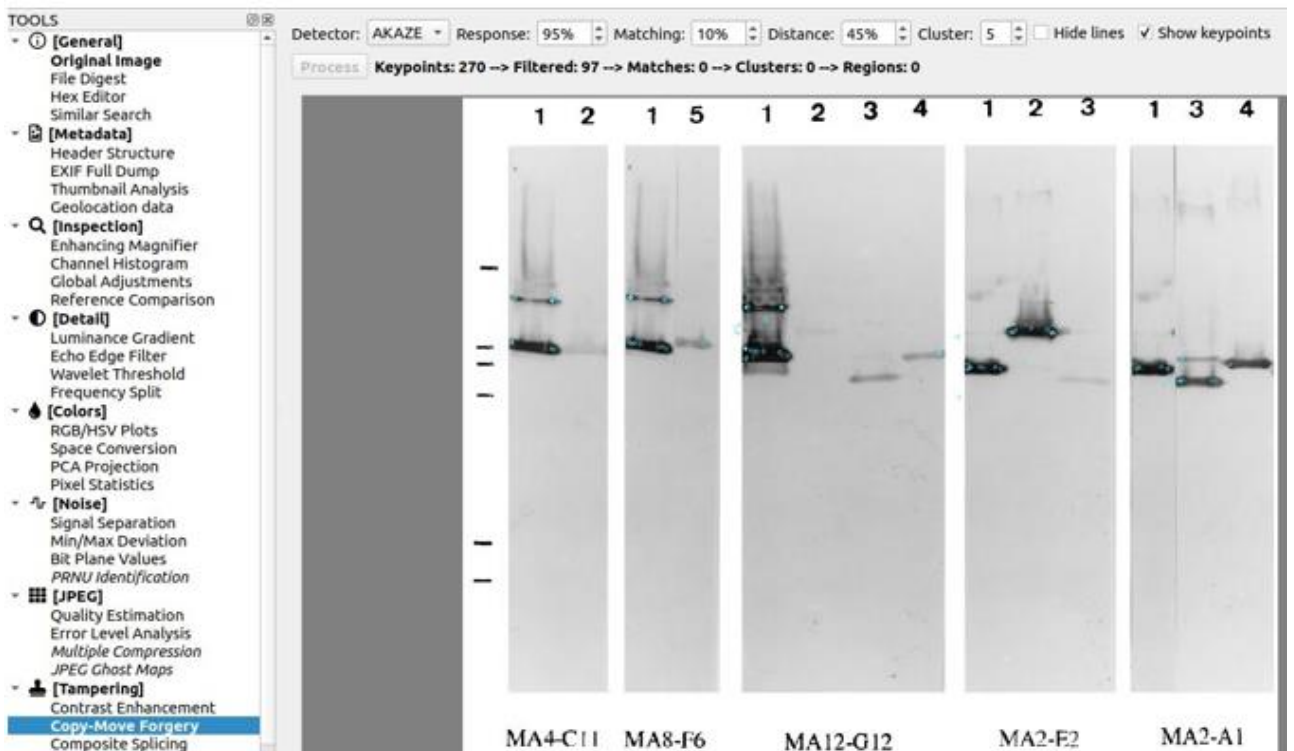

With a sensitivity threshold of 10%, we find results that could be similar to those challenged by BIK.

Finally, with the AKAZE (Accelerated KAZE) method, **no copies are detected**.

4. **Analysis of one of Bik's publication** - https://journals.asm.org/doi/10.1128/mbio.00809-16

Our analysis and methods would not be complete if we did not apply them to **a search for duplicates** in a study published by Bik.

Using the same methodology, we took a look at image 4 from her publication, and more specifically at zones 9 and 10. At 0.999 threshold , **no duplicates are identified**, whereas at 0.96, one zone is identified as duplicated!

**BIK'S PUBLICATION ANALYSES – TEST ON IMAGE 4** France-Soir
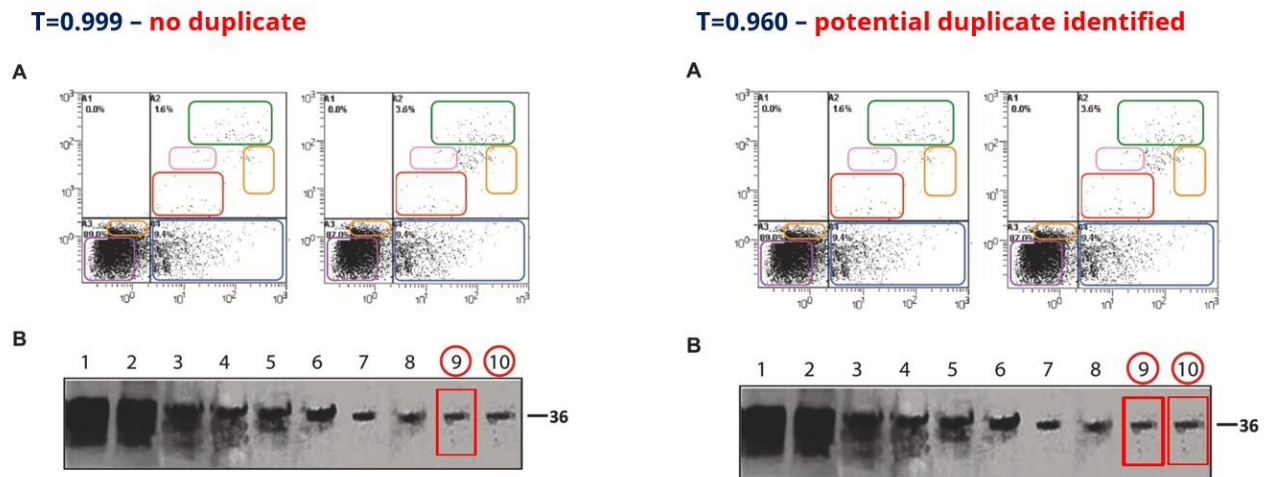
In image 2, **no duplicate at the 0.999 threshold**, but a duplicate at 0.99. This **could be a duplicate image, as the difference is very small indeed**. However, this is a Bik publication, so it would be churlish to make such a claim. Unless this is a demonstration that "charity begins at home", and that Bik should have checked his methods on his own images.
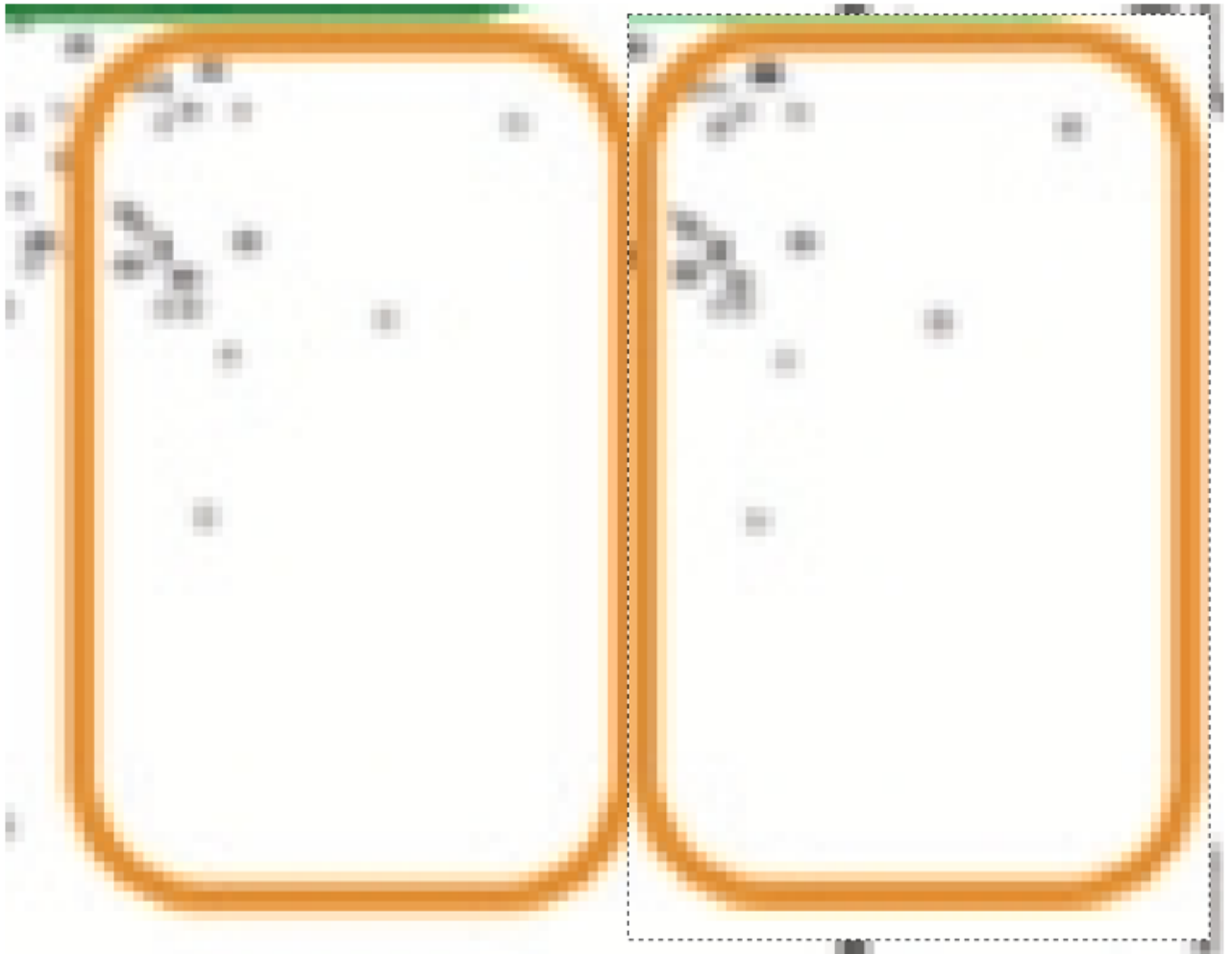
BIK'S PUBLICATION ANALYSES – TEST ON IMAGE 2 France-Soir
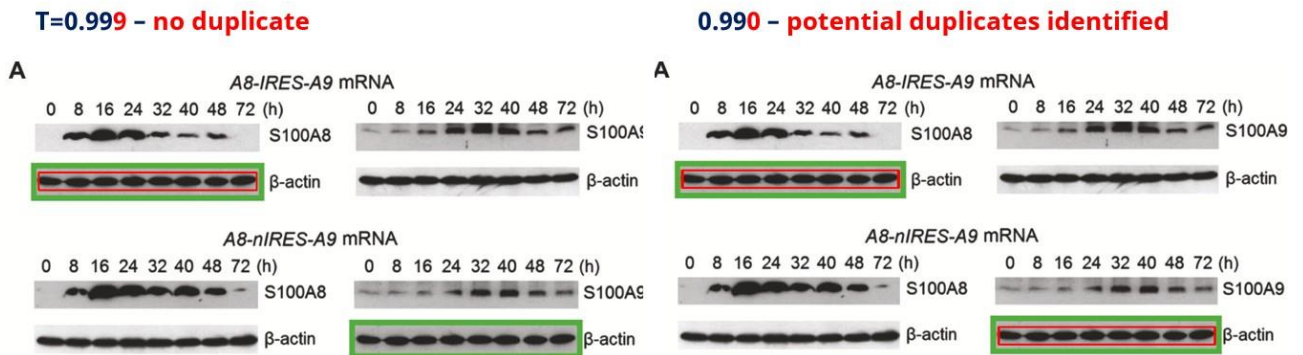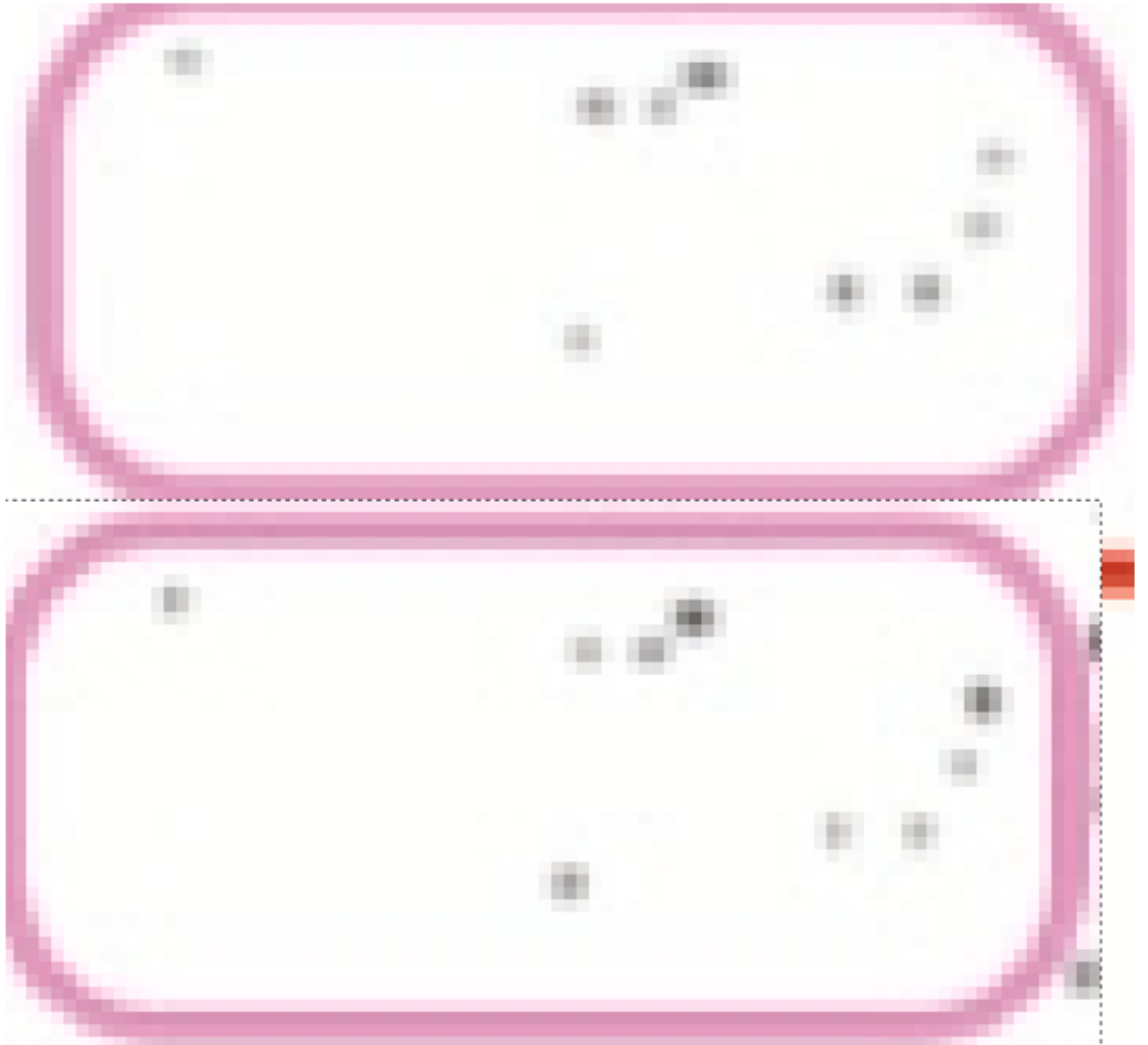

Let's take the analysis a step further and look at the different-colored areas in image 4 below:
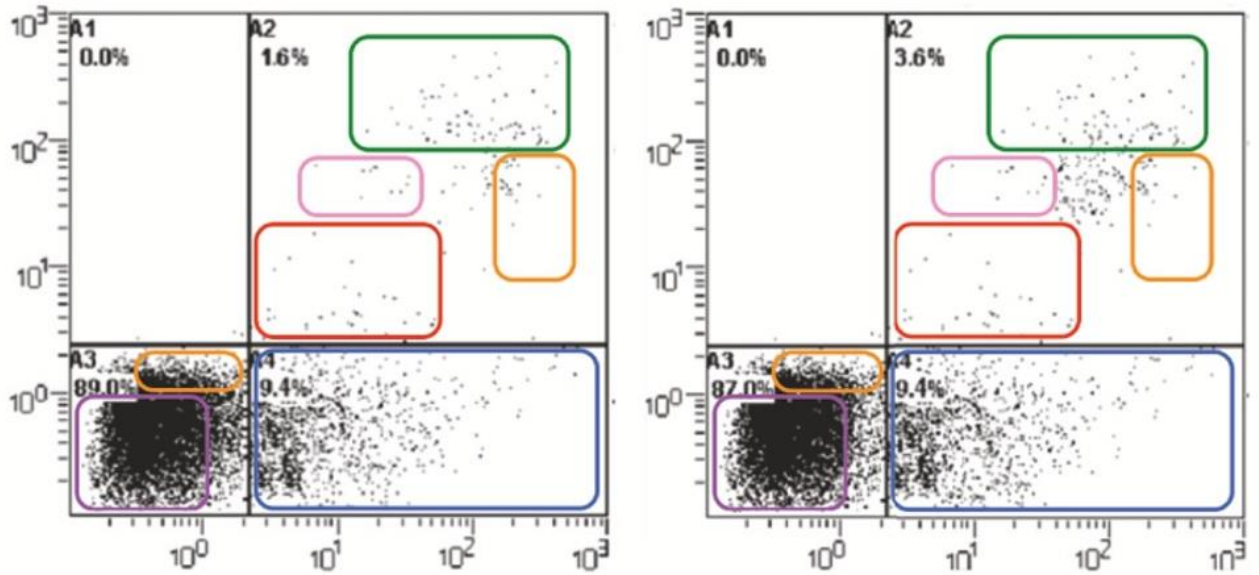
The difference is visible to the naked eye **by simply by placing the zones side by side**.

In general, for image verification analysis, we recommend the use of Sherloq, an open source software. Bik talks about a site without identifying whether it's the same one. https://29a.ch/photo-forensics/#clone-detection

Let's check this with Nasa data from the James Webb telescope: with the FotoForensics site, numerous clones/markers can be detected.

JAMES WEB IMAGE VERIFICATION WITH PHOTO FORENSICS
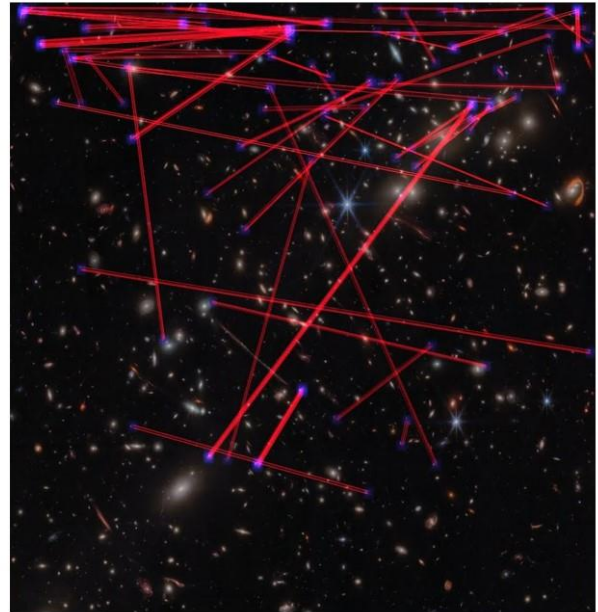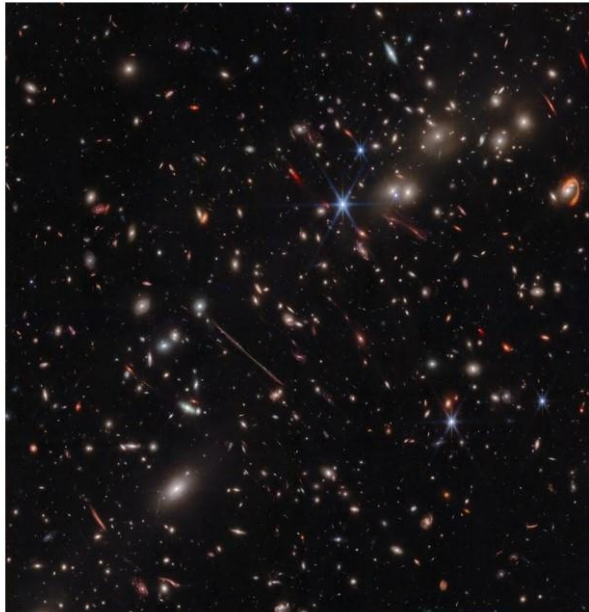
But they are not confirmed with Sherloq !
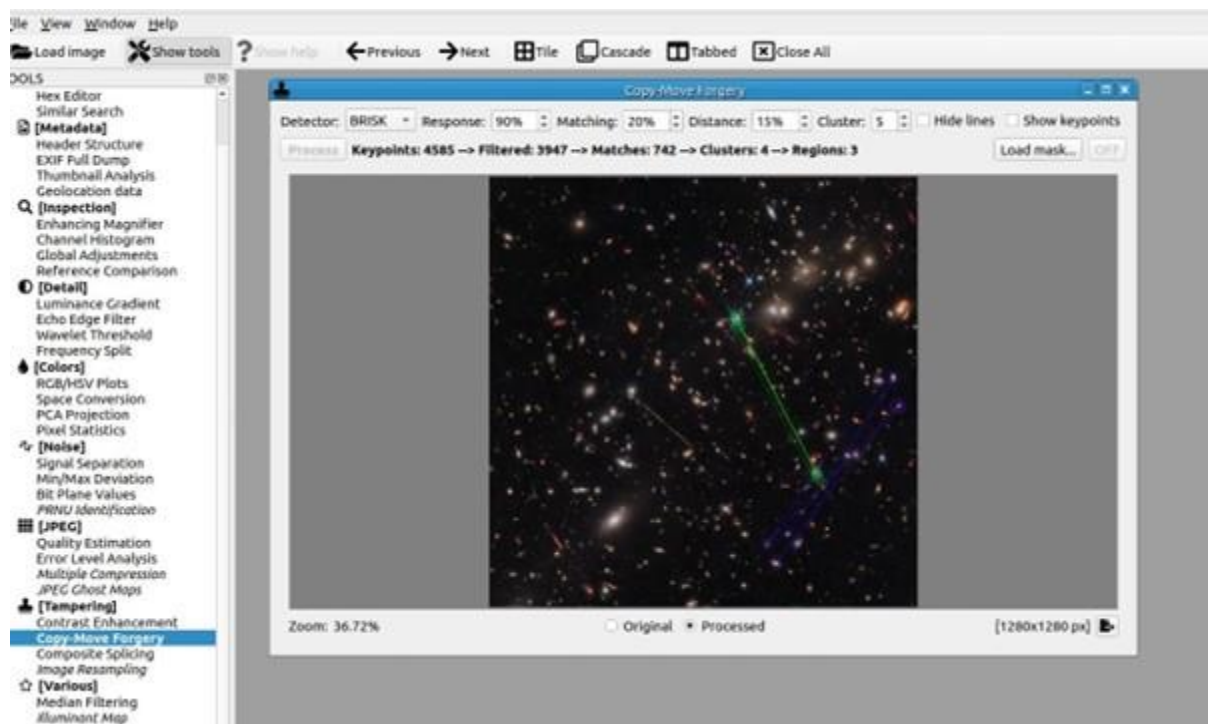
Conclusions

After verifying Bik's analyses, we have demonstrated that :

- Her claims are **unfounded**.
- Her error(s) is/are due to an underestimation of 3% difference.
- The errors are not due to the JPEG compression of the image in the PDF file.
- Two methodologies have been applied to be sure of the result (although none is indicated, apart from Software as Service).
- The same conclusions have been demonstrated for other publications.
- In addition, we have shown that she does not use the software she herself advises to use on her blog (Forensically).

Her claimed intention to **expose scientific fraud is laudable**, as scientific integrity is the basis of trust in science. However, **one can only wonder about her intentions when she attacks IHU publications**. And who pays her to attack the work of researchers with a pseudo-scientific approach that cannot be replicated?

Moreover by using her own fraud-hunting methods on her own publication, the latter is subject to the same criticism. The question remains: why doesn't she verify her method on her own publications before applying it to others?

Bik's **conclusions are therefore unfounded**, and **she should withdraw her comments** at the risk of being called a fraud herself. The biter being bit.

6 February 2024