Evolutionary Genetics

LV 25600-01 | Lecture with exercises | 4KP



Jean-Claude Walser

jean-claude.walser [at] env.ethz.ch











Erfundene Tierversuche in Labor des Unispitals

Falsche Mäusegehirne bringen Starforscher in Bedrängnis

Adriano Aguzzi ist einer der bedeutendsten Wissenschaftler der Schweiz. Doch nun wackelt das Denkmal des Neuropathologen: An seinem Institut wurden Forschungsergebnisse gefälscht – und weder er noch die Universität Zürich haben transparent über die Vorfälle informiert.

False mouse brains put star researchers in trouble

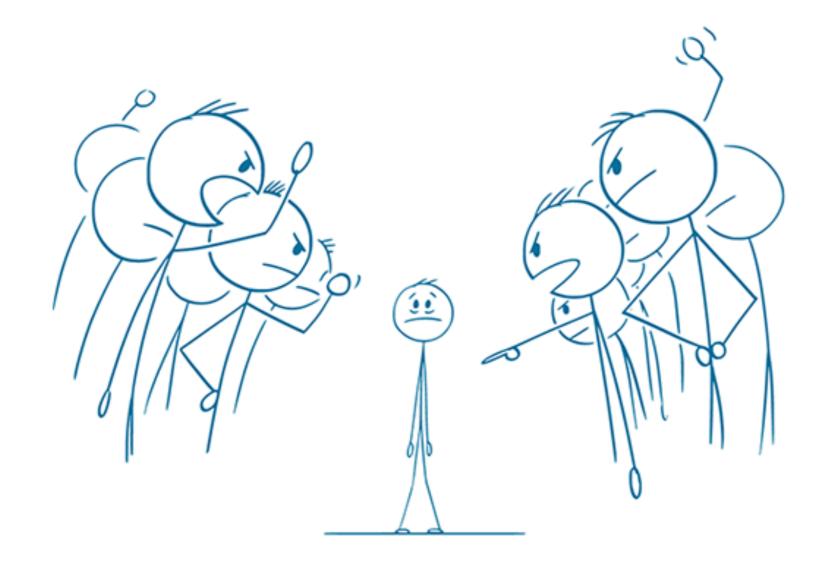


Front Wetter US-Wahlen Nahostkonflikt Schweiz #WIRSINDZUKUNFT Sport Zürich Ber

Gefälschte Tierversuche setzen Star-Forscher unter Druck

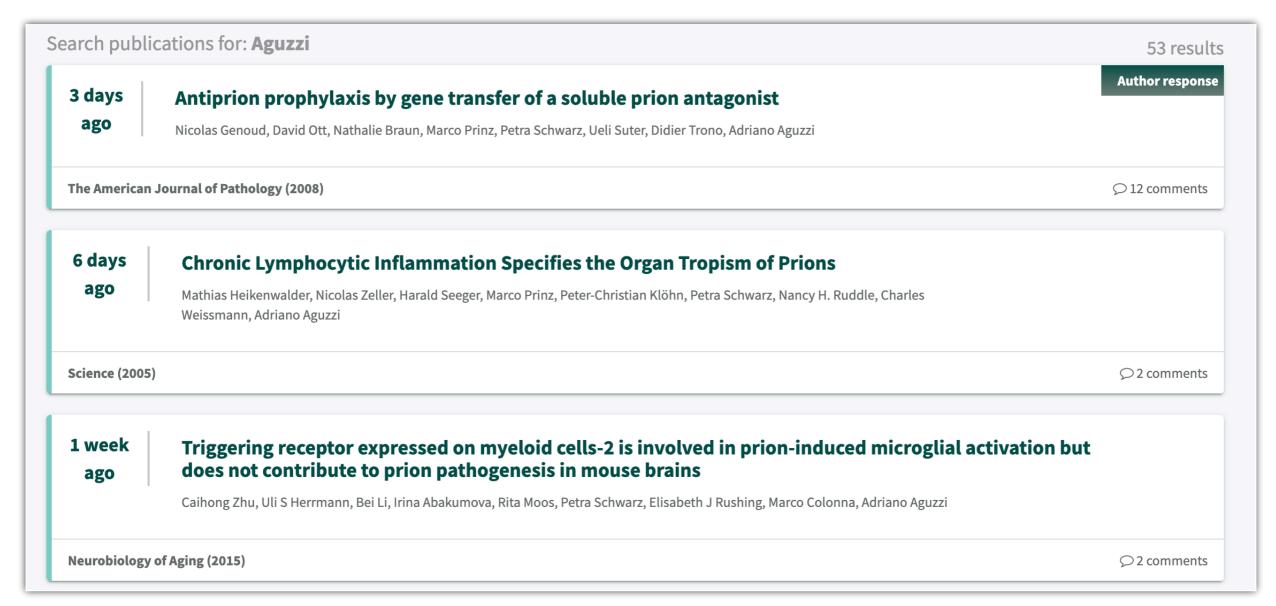
Ein ehemaliger Mitarbeiter von Adriano Aguzzi am Institut für Neuropathologie am Unispital Zürich manipulierte Daten von Tierversuchen.

How can we be sure that it is not a scapegoat?

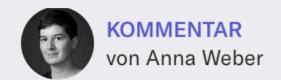


scapegoat - a person who is blamed for the wrongdoings, mistakes, or faults of others, especially for reasons of expediency.



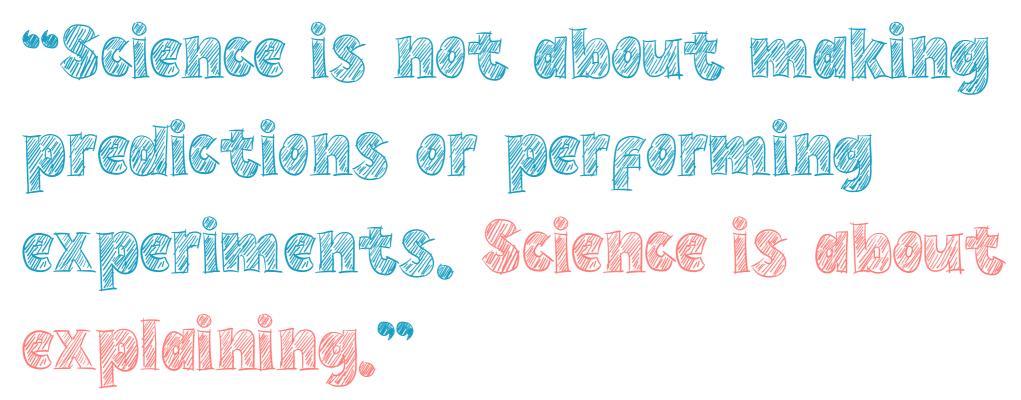


NZZ



Fälschungsskandal an der Uni Zürich: Der Fall Adriano Aguzzi steht für längst bekannte Probleme in der Forschung





→ Bill Gaede

^{*}Guillermo "Bill" Gaede is an Argentine engineer and programmer who is best known for Cold War industrial spying conducted while he worked at AMD and Intel.

Title **Abstract** Introduction **Materials and Methods** Results **Discussion** Conclusion

Data Availability **Author Contributions** Funding

Acknowledgments

Supplementary Material

References

The mutational spectrum of non-CpG DNA varies with CpG content

Jean-Claude Walser¹ and Anthony V. Furano²

Section on Genomic Structure and Function, Laboratory of Malecular and Cellular Biology, National Institute of Diabetes and Digestive and Kidney diseases, National Institutes of Health, Bethesda, Maryland 20892-0830, USA

The accumulation of base substitutions (mutations) not subject to natural selection is the neutral mutation rate. Because this rate reflects the in vivo processes involved in maintaining the integrity of genetic information, the factors that affect the neutral mutation rate are of considerable interest. Mammals exhibit two dramatically different neutral mutation rates the CpG mutation rate, wherein the C of most CpGs (Le., methyl-CpG) mutate at 10-50 times that of C in any other context or of any other base. The latter mutations constitute the non-CpG rate. The high CpG rate results from the spontaneous deamination of methyl-C to T and incomplete restoration of the ensuing T:G mismatches to C:Gs. Here, we determined the neutral non-CpG mutation rate as a function of CpG content by comparing sequence divergence of thousands of pairs of neutrally evolving chimpanaee and human orthologs that differ primarily in CpG content. Both the mutation rate and the mutational spectrum (transition/transversion ratio) of non-CpG residues change in parallel as sigmoidal (logistic) functions of CpG content. As different mechanisms generate transitions and transversions, these results indicate that both mutation rate and mutational processes are contingent on the local CpG content. We consider several possible mechanisms that might explain how CpG exerts these effects.

[Supplemental material is available online at http://www.genome.org.]

DNA base substitutions (mutations) are the most frequent class of genetic variants. Thus, determining the factors that affect the base mutation rate (i.e., the number of base substitutions over time) remains a major concern of geneticists and molecular evolutionists (e.g., Nachman and Crowell 2000; Hwang and Green 2004; Duret 2009). Mutations not subject to natural selection are considered neutral and the neutral mutation rate is considered to closely reflect or equal the actual mutation sate (Ochman 2003). Thus, the neutral mutation rate is a basic biological parameter, which can be estimated from the number of interspecies base differences (sequence divergence) be-tween neutrally evolving orthologous sequences (i.e., those sharing

a common ancestral sequence, e.g., Nachman and Crowell 2000).

The neutral mutation rate varies considerably between and within chromosomes. Although numerous factors have been conrelated with the neutral mutation rate (e.g., Krawczak et al. 1998; Hardhon et al. 2003; Hwang and Green 2004; Chimpanzee Sequencing Analysis Consortium 2005; Gaffney and Keightley 2005; Hellmann et al. 2005; Taylor et al. 2006), the mechanism(s) accounting for these correlations remain elusive (e.g., Hodgkinson et al. 2009; for review, see Duret 2009).

One of the more intriguing covariates of the neutral mutation sate is its positive correlation with CpG content. In part, this corre-lation is not surprising because most CpGs in mammals are uniquely hypermutable (e.g., 19vang and Green 2004). The Cs of most CpCs are methylated (Ehrlich and Wang 1981), which enhances the deamination of C, in this case producing a TG mismatch. The net result is that methyl-CpGs mutate at 10-50 times the rate of C in any other context (Coulondre et al. 1978; Duncan and Miller 1980; Bulmer 1986: Seed and Bird 1990), or of any other base (Hwang and

Green 2004). Consequently, CpGs not under selection are replaced

over time by TpG/CpAs.
Inexplicably, however, the positive correlation between CpG content and the neutral mutation rate persists even if mutation at CpG sites are not counted, i.e., if only non-CpG mutations are measured (Chimpanzee Sequencing Analysis Consortium 2005; Gaffney and Keightley 2005; Hellmann et al. 2005; Tyekucheya et al. 2008). A prevailing reasonable explanation for this odd result was that the non-CpG mutation rate and CpG content were joint manifestations of the chromosomal environment (Hellmann et al

We recently considered an alternative, that CpGs (i.e., methyl CpGs), or mutations thereof, somehow directly affect the mutation of flanking non-CpG DNA. This explanation would have far-reaching implications given the epigenetic role of CpG methylation in gene regulation, chromatin structure, imprinting, and the tilencing of transposable elements and other genomic insertions (Lees-Murdock and Waish 2008; Cedar and Bergman 2009). We addressed this issue by examining the sequence divergence of thousands of neutrally evolving orthologous sequences in the chimpanase and human ge-

nomes that differed primarily in CpG content (Walser et al. 2008).

These orthologs were the repeated DNA fossils that had been interspersed throughout the genome at different times in the pri mate lineage of humans and chimpaneres by six now extinct fami-lies of L1 non-LTR retrotransposons. As L1 fossis are not under se lection, the CpG content of these otherwise very similar sequences should differ. Thus, the CpG content of the younger L1 fossils should be higher than that of the older ones and, if our supposition was correct, so should their mutation rate, regardless of chromosomal location. And this is the result we obtained (Walser et al. 2008).

Our current examination of non-CpG mutations over a considerably wider range of CpG content than previously (Waise et al. 2008) revealed two unexpected findings: First, the correlation between the two is best fit by a sigmoid (logistic) function. Thus, both a certain threshold CpG content is required to have a substan-tial effect on the overall non-CpG mutation rate, and "saturation"

20:000-000; ISSN 1088-9051/10; www.genome.org

Genome Research 1

Title

Abstract

Introduction

Materials and Methods

Results

Discussion

Conclusion

Data Availability

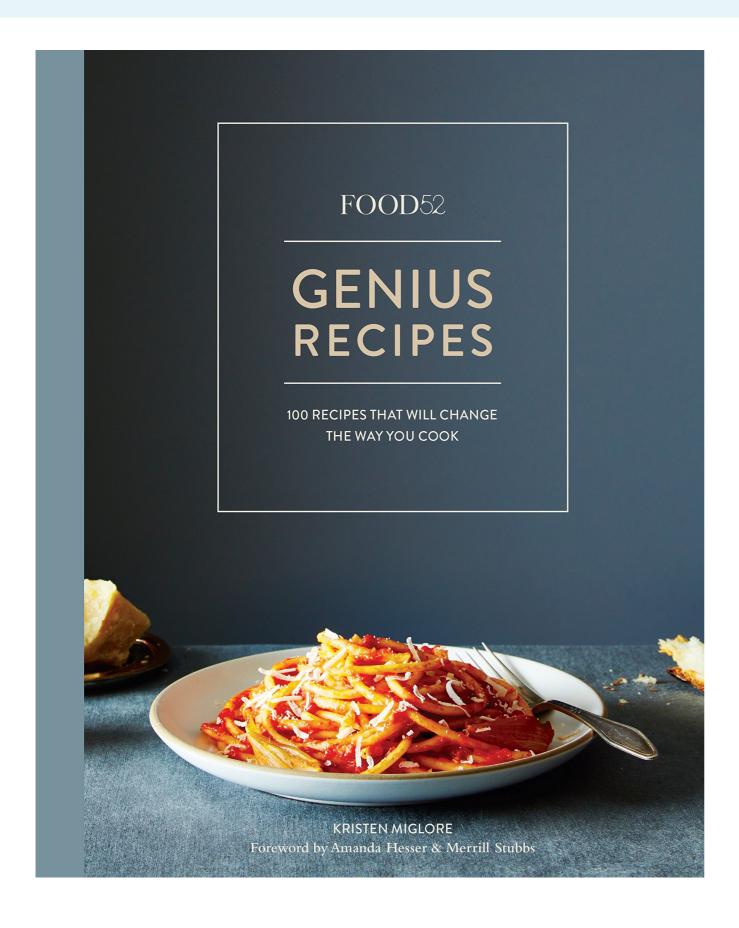
Author Contributions

Funding

Acknowledgments

Supplementary Material

References



Roasted Applesauce

FROM JUDY RODGERS

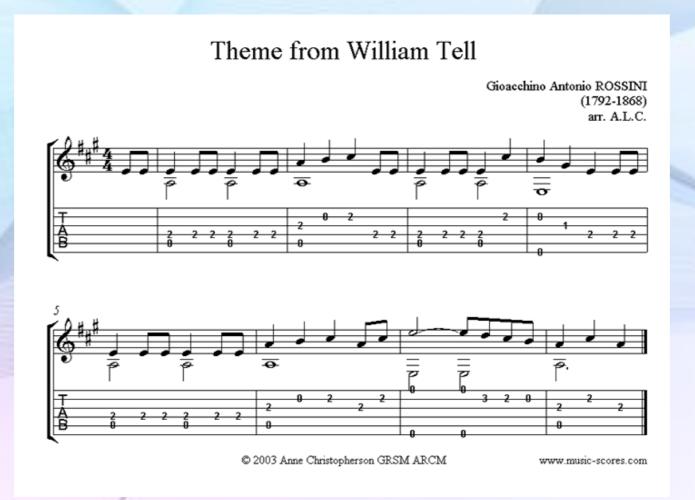
No cinnamon, no cloves—this sauce is straight apple.

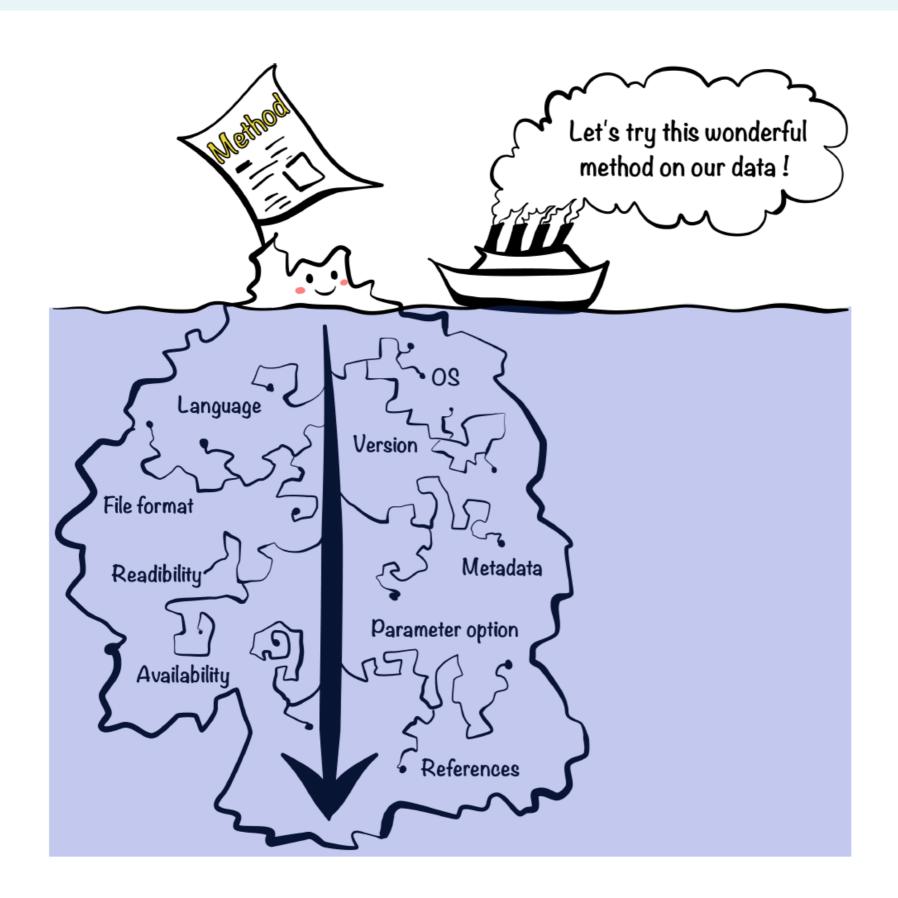
It comes from Judy Rodgers's Zuni Café Cookbook and—as with everything served at her San Francisco restaurant—it's smart and simple, balancing the apples only as needed with small amounts of salt, sugar, and apple cider vinegar. Not only does this quick oven method free you from stewing and stewing applesauce on the stovetop, but it does the magic that roasting always does. All the sugars concentrate, allowing apples to become the best version of themselves. There's just a little bit of butter too, sliced into wafers that melt into bronzed apple tops and a rich sauce.

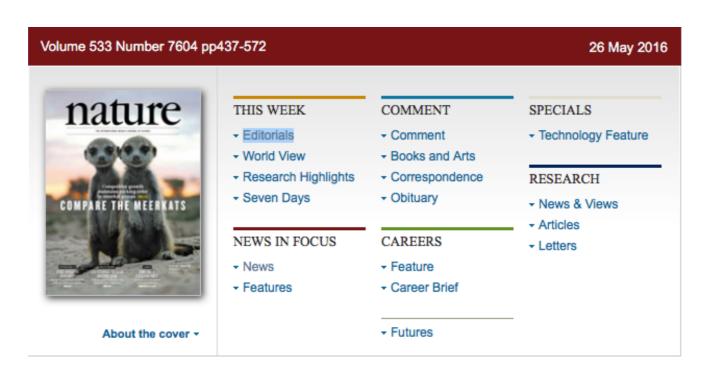
Makes about 3 cups (710ml)

3½ to 4 pounds (1.6 to 1.8kg) apples (Rodgers recommends crisp eating apples, like Sierra Beauties, Braeburns, Pippins, Golden Delicious, or Galas)
Pinch of salt
Up to 2 teaspoons sugar, as needed
About 2 tablespoons unsalted butter
A splash of apple cider vinegar, as needed

- 1 Preheat oven to 375°F (190°C).
- 2 Peel, core, and quarter the apples. Toss with a little salt and, unless they are very sweet, a bit of sugar to taste. If they are tart enough to make you squint, add the full measure of sugar. Spread in a shallow baking dish that crowds the apples in a single layer. Drape with slivers of the butter, cover tightly with a lid or aluminum foil, and bake until the apples start to soften, 15 to 30 minutes, depending on your apples.
- 3 Uncover, raise the heat to 500°F (260°C), and return the pan to the oven. Leave the apples to dry out and color slightly, about 10 minutes.
- 4 When the tips of the apples have become golden and the fruit is tender, scrape them into a bowl and stir into a chunky "mash." Season with salt and sugar to taste, then consider a splash of apple cider vinegar to brighten the flavor. (Try a drop on a spoonful to see if you like it.) Serve hot or warm.

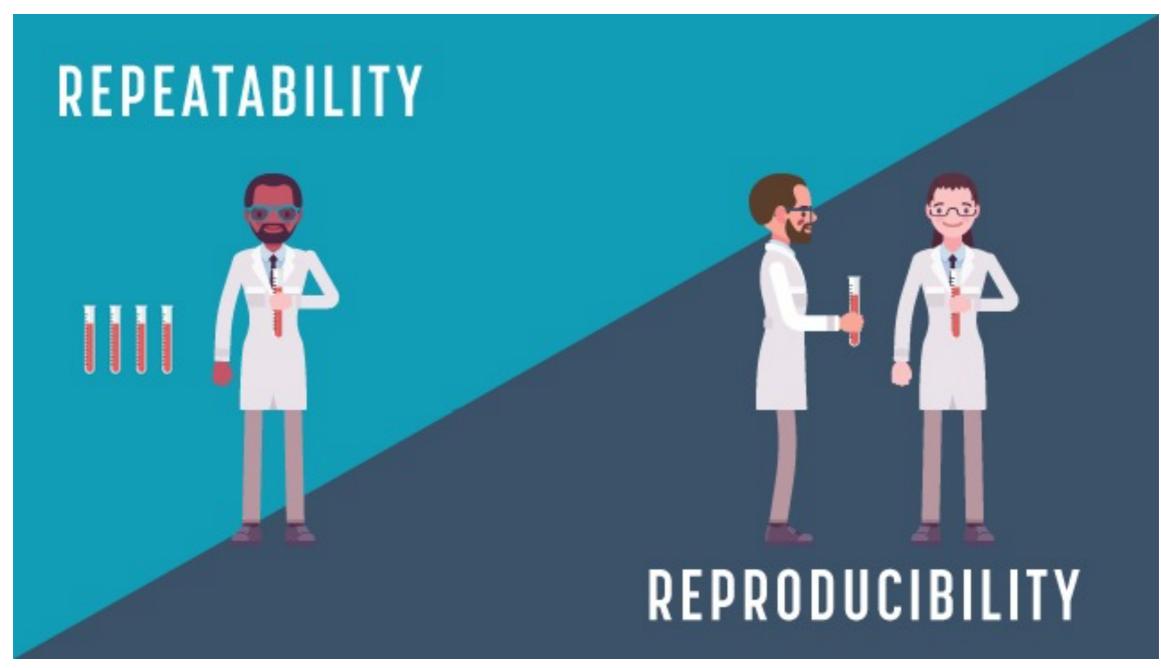






Reality check on reproducibility

A survey of Nature readers revealed a high level of concern about the problem of irreproducible results. Researchers, funders and journals need to work together to make research more reliable.



Source: https://www.technologynetworks.com/informatics/articles/repeatability-vs-reproducibility

Repeatability is a measure of the likelihood that, having produced one result from an experiment, you can try the same experiment, with the same setup, and produce that same result. It is a way for researchers to verify that their own results are true and are not just **chance artefacts**.

The **reproducibility** of data is a measure of whether a different research team can attain results published in a paper using the same methods. This shows that the results are **not artefacts of the unique setup in one research lab**. It is easy to see why reproducibility is desirable, as it reinforces findings and protects against rare cases of fraud, or less rare cases of human error, in the production of significant results.

Replicability - Different team, different experimental setup. If an observation is replicable it should be able to be made by a different team, using a different measuring system and dataset, in a different location, on multiple trials. This would therefore involve collecting data anew.

Source: https://www.technologynetworks.com/informatics/articles/repeatability-vs-reproducibility

1 How can we improve repeatability?

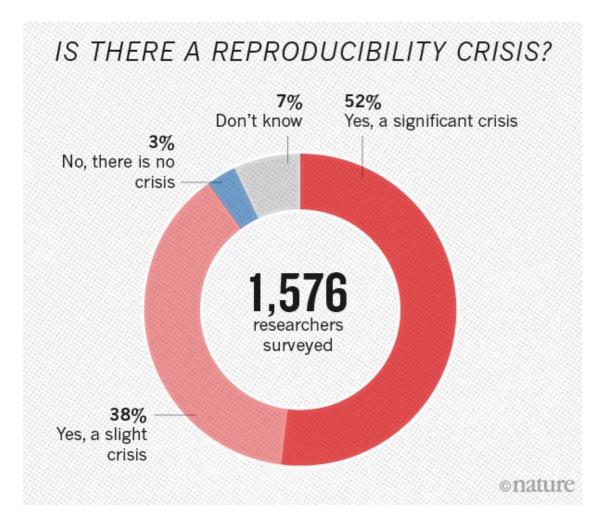
O2 How can we improve reproducibility?

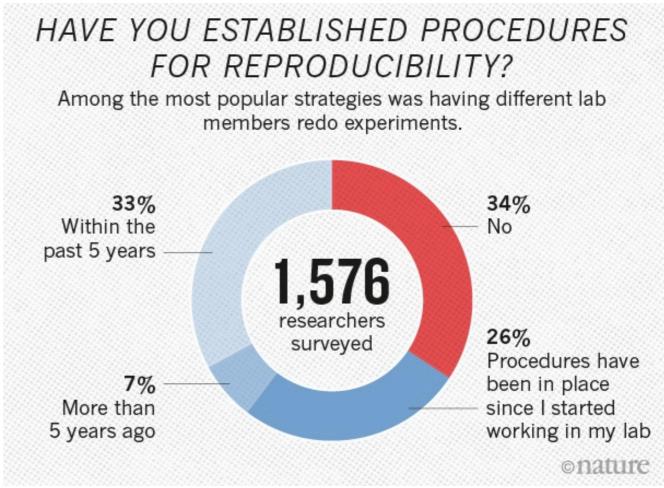
O3 How can we improve replicability?



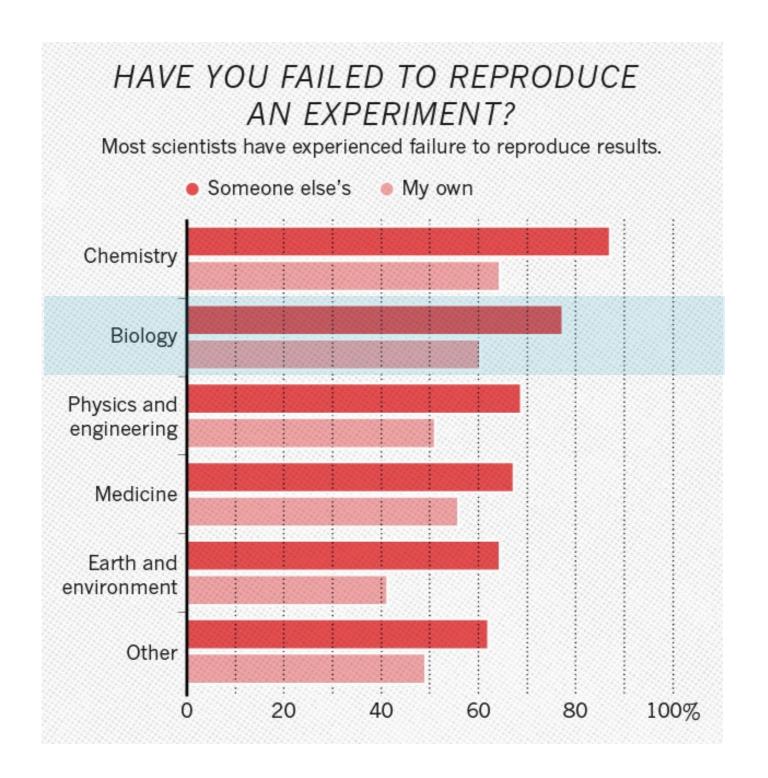
YouTube: Is there a reproducibility crisis in science? - Matt Anticole

The collective effort of science depends on researchers being able to reproduce the work of others. In a recent survey of 1,576 researchers, 70% of them admitted having difficulty in reproducing experiments proposed by other scientists. For 50%, this reproducibility issue even concerns their own experiments.

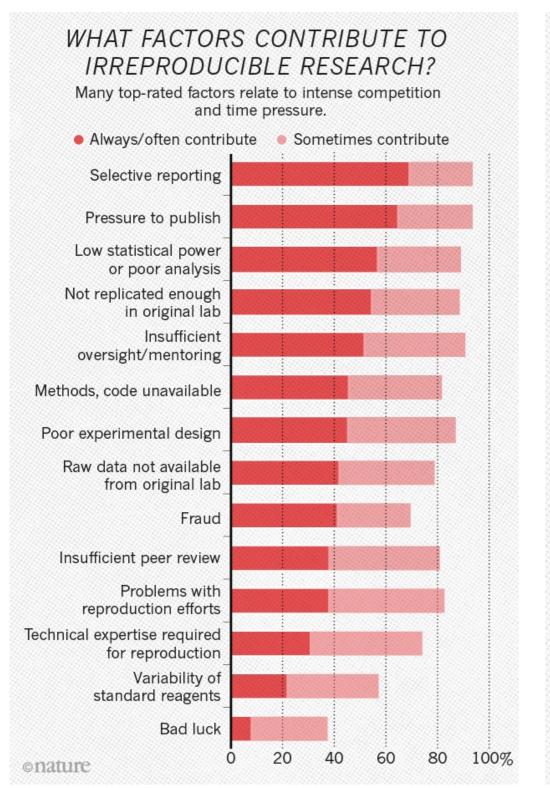


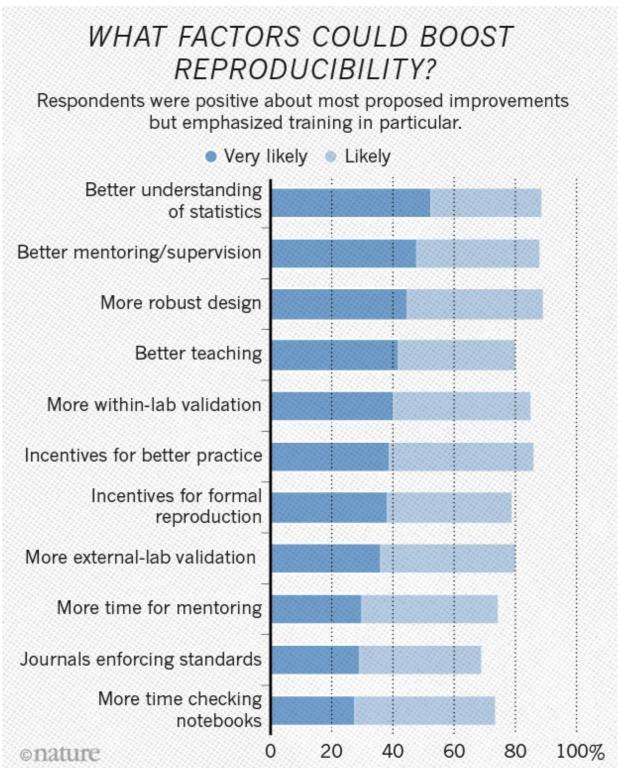


Source: Baker (2016) 1,500 scientists lift the lid on reproducibility. Nat News . 533:452.



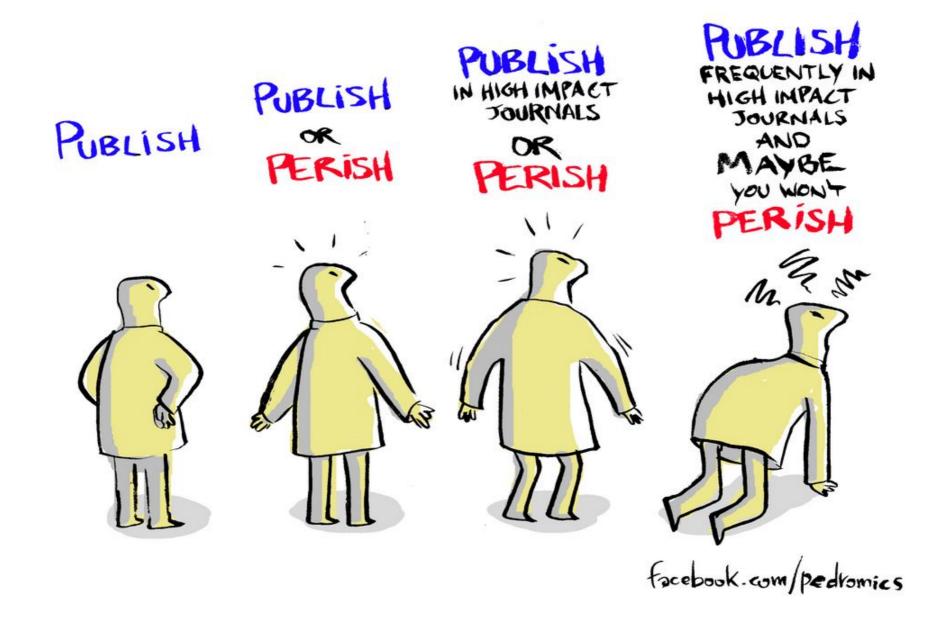
Source: Baker (2016) 1,500 scientists lift the lid on reproducibility. Nat News . 533:452.





Source: Baker (2016) 1,500 scientists lift the lid on reproducibility. Nat News . 533:452.

THE EVOLUTION OF ACADEMIA





- Honest Mistakes
- Careless Mistakes
- Cheats / Fraud

Molecular Breeding January 2015, 35:54

Date: 25 Jan 2015

Development of a leafy *Brassica rapa* fixed line collection for genetic diversity and population structure analysis

Wenxing Pang, Xiaonan Li, Su Ryun Choi, Vignesh Dhandapani, Subin Im, Min Young Park, Chang Soon Jang, Man-Sung Yang, In Ki Ham, Eun Mo Lee, Wankyu Kim, Soo-Seong Lee, Guusje Bonnema, Suhyoung Park, Zhongyun Piao, Yong Pyo Lim



Article Metrics

(Illumina). Paired-end short read sequences for Kenshin were generated by the Illumina Genome Analyzer-IIx system. Low quality Illumina reads were identified and trimmed using fastqc (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The Bowtie2 aligner was used to align the high-quality short read sequences of Kenshin to the reference genome sequence and a BAM alignment file was generated along with the consensus sequence. SAMtools and BCFtools were employed to call SNPs and

FastQC is for quality control but not for data manipulations.

An Apple a Day: Which Bacteria Do We Eat With Organic and Conventional Apples?



ORIGINAL RESEARCH

published: 24 July 2019 doi: 10.3389/fmicb.2019.01629

Birgit Wassermann, Henry Müller and Gabriele Berg*

Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

Apples are among the most consumed fruits world-wide. They represent a source of direct human exposure to bacterial communities, which is less studied. We analyzed the apple microbiome to detect differences between tissues and the impact of organic and conventional management by a combined approach of 16S rRNA gene amplicon analysis and qPCR, and visualization using fluorescence in situ hybridization and confocal laser scanning microscopy (FISH-CLSM). Each apple fruit harbors different tissues (stem, peel, fruit pulp, seeds, and calyx), which were colonized by distinct bacterial communities. Interestingly, fruit pulp and seeds were bacterial hot spots, while the peel was less colonized. In all, approximately 108 16S rRNA bacterial gene copy numbers were determined in each g apple. Abundances were not influenced by the management practice but we found a strong reduction in bacterial diversity and evenness in conventionally managed apples. In addition, despite the similar structure in general dominated by Proteobacteria (80%), Bacteroidetes (9%), Actinobacteria (5%), and Firmicutes (3%), significant shifts of almost 40% of bacterial genera and orders were monitored. Among them, especially bacterial signatures known for health-affecting potential were found to be enhanced in conventionally managed apples. Our results suggest that we consume about 100 million bacterial cells with one apple. Although this amount was the same, the bacterial composition was significantly different in conventionally and organically produced apples.

What is the health-relevant information from the abstract and can you spot a problem?

An Apple a Day: Which Bacteria Do We Eat With Organic and Conventional Apples?

Birgit Wassermann, Henry Müller and Gabriele Berg*

Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

Apples are among the most consumed fruits world-wide. They represent a source of direct human exposure to bacterial communities, which is less studied. We analyzed the apple microbiome to detect differences between tissues and the impact of organic and conventional management by a combined approach of 16S rRNA gene amplicon analysis and qPCR, and visualization using fluorescence in situ hybridization and confocal laser scanning microscopy (FISH-CLSM). Each apple fruit harbors different tissues (stem, peel, fruit pulp, seeds, and calyx), which were colonized by distinct bacterial communities. Interestingly, fruit pulp and seeds were bacterial hot spots, while the peel was less colonized. In all, approximately 108 16S rRNA bacterial gene copy numbers were determined in each g apple. Abundances were not influenced by the management practice but we found a strong reduction in bacterial diversity and evenness in conventionally managed apples. In addition, despite the similar structure in general dominated by Proteobacteria (80%), Bacteroidetes (9%), Actinobacteria (5%), and Firmicutes (3%), significant shifts of almost 40% of bacterial genera and orders were monitored. Among them, especially bacterial signatures known for health-affecting potential were found to be enhanced in conventionally managed apples. Our results suggest that we consume about 100 million bacterial cells with one apple. Although this amount was the same, the bacterial composition was significantly different in conventionally and organically produced apples.



In all, approximately 10⁸ 16S rRNA bacterial **gene copy numbers** were determined in each **g apple**.

Our results suggest that we consume about **100 million** bacterial **cells** with one **apple**.

An Apple a Day: Which Bacteria Do We Eat With Organic and Conventional Apples?

Birgit Wassermann, Henry Müller and Gabriele Berg*

Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

Apples are among the most consumed fruits world-wide. They represent a source of direct human exposure to bacterial communities, which is less studied. We analyzed the apple microbiome to detect differences between tissues and the impact of organic and conventional management by a combined approach of 16S rRNA gene amplicon analysis and qPCR, and visualization using fluorescence in situ hybridization and confocal laser scanning microscopy (FISH-CLSM). Each apple fruit harbors different tissues (stem, peel, fruit pulp, seeds, and calyx), which were colonized by distinct bacterial communities. Interestingly, fruit pulp and seeds were bacterial hot spots, while the peel was less colonized. In all, approximately 108 16S rRNA bacterial gene copy numbers were determined in each g apple. Abundances were not influenced by the management practice but we found a strong reduction in bacterial diversity and evenness in conventionally managed apples. In addition, despite the similar structure in general dominated by Proteobacteria (80%), Bacteroidetes (9%), Actinobacteria (5%), and Firmicutes (3%), significant shifts of almost 40% of bacterial genera and orders were monitored. Among them, especially bacterial signatures known for health-affecting potential were found to be enhanced in conventionally managed apples. Our results suggest that we consume about 100 million bacterial cells with one apple. Although this amount was the same, the bacterial composition was significantly different in conventionally and organically produced apples.

Apple are a "source" of bacteria. What a big suprise!

Fruit pulp and seeds are bacterial hot spots. Think again!

What? Good, bad or what? Outside or inside an apple?

"Read not to contradict and confute; nor to believe and take for granted; nor to find talk and discourse; but to weigh and consider."

—Francis Bacon

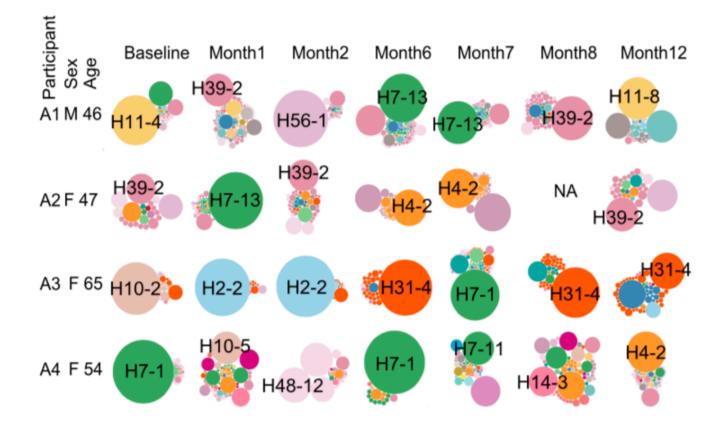
Cell Reports



Resource

Single-gene long-read sequencing illuminates *Escherichia coli* strain dynamics in the human intestinal microbiome

Dalong Hu,¹ Nicholas R. Fuller,^{2,3} Ian D. Caterson,^{1,2,3} Andrew J. Holmes,¹ and Peter R. Reeves^{1,4,*}



Flagellin diversity was discovered as strain differences in im- mune responses, which were codified and used for serotyping. The variable flagellin domains are known in serology as the H an- tigen, and 53 H antigens were distinguished in E. coli, of which 52 have the high-level divergence, with essentially no significant sequence alignment

An Apple a Day: Which Bacteria Do We Eat With Organic and Conventional Apples?

frontiers in Microbiology

ORIGINAL RESEARCH

published: 24 July 2019 doi: 10.3389/fmicb.2019.01629

Birgit Wassermann, Henry Müller and Gabriele Berg*

Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

Apples are among the most consumed fruits world-wide. They represent a source of direct human exposure to bacterial communities, which is less studied. We analyzed the apple microbiome to detect differences between tissues and the impact of organic and conventional management by a combined approach of 16S rRNA gene amplicon analysis and qPCR, and visualization using fluorescence in situ hybridization and confocal laser scanning microscopy (FISH-CLSM). Each apple fruit harbors different tissues (stem, peel, fruit pulp, seeds, and calyx), which were colonized by distinct bacterial communities. Interestingly, fruit pulp and seeds were bacterial hot spots, while the peel was less colonized. In all, approximately 108 16S rRNA bacterial gene copy numbers were determined in each g apple. Abundances were not influenced by the management practice but we found a strong reduction in bacterial diversity and evenness in conventionally managed apples. In addition, despite the similar structure in general dominated by Proteobacteria (80%), Bacteroidetes (9%), Actinobacteria (5%), and Firmicutes (3%), significant shifts of almost 40% of bacterial genera and orders were monitored. Among them, especially bacterial signatures known for health-affecting potential were found to be enhanced in conventionally managed apples. Our results suggest that we consume about 100 million bacterial cells with one apple. Although this amount was the same, the bacterial composition was significantly different in conventionally and organically produced apples.

This is a typical study of a very limited and questionable design, were the authors made many mistakes, did not include any controls, group whatever they liked, interpreted the reults according to their personal belives, and did not care about reproducibility (for good reasons I guess?).

Science Integrity Digest

A blog about science integrity, by Elisabeth Bik, for Harbers-Bik LLC. Support my work at Patreon.com/elisabethbik



eliesbik

February 15, 2024

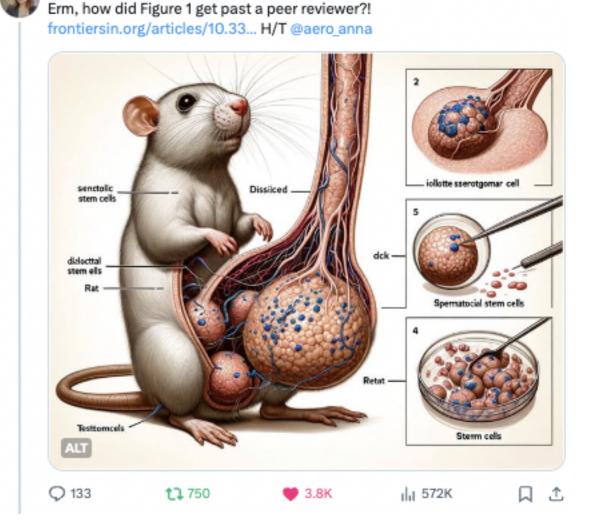
The rat with the big balls and the enormous penis – how Frontiers published a paper with botched Algenerated images

A review article with some obviously fake and non-scientific illustrations created by Artificial Intelligence (AI) was the talk on X (Twitter) today.

The figures in the paper were generated by the AI tool Midjourney, which generated some pretty, but nonsensical, illustrations with unreadable text.

It appears that neither the editor nor the two peer reviewers looked at the figures at all. The paper was peerreviewed within a couple of weeks and published two days ago.

Dear readers, today I present you: the rat with the enormous family jewels and the dislocttal stem ells.



Dr CJ Houldcroft * @DrCJ_Houldcroft · 3h



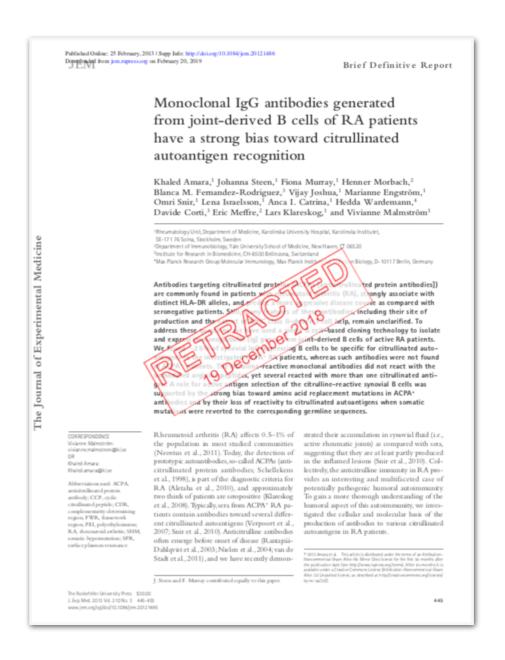
A database search with the query affiliation "Zurich" and country "Switzerland" showed **73 hits**.

A database search with the query affiliation "ETH" and country "Switzerland" showed **27 hits**.

A database search with the query affiliation "University of Zurich" and country "Switzerland" showed **16 hits**.

https://retractionwatch.com

http://retractiondatabase.org/RetractionSearch.aspx?



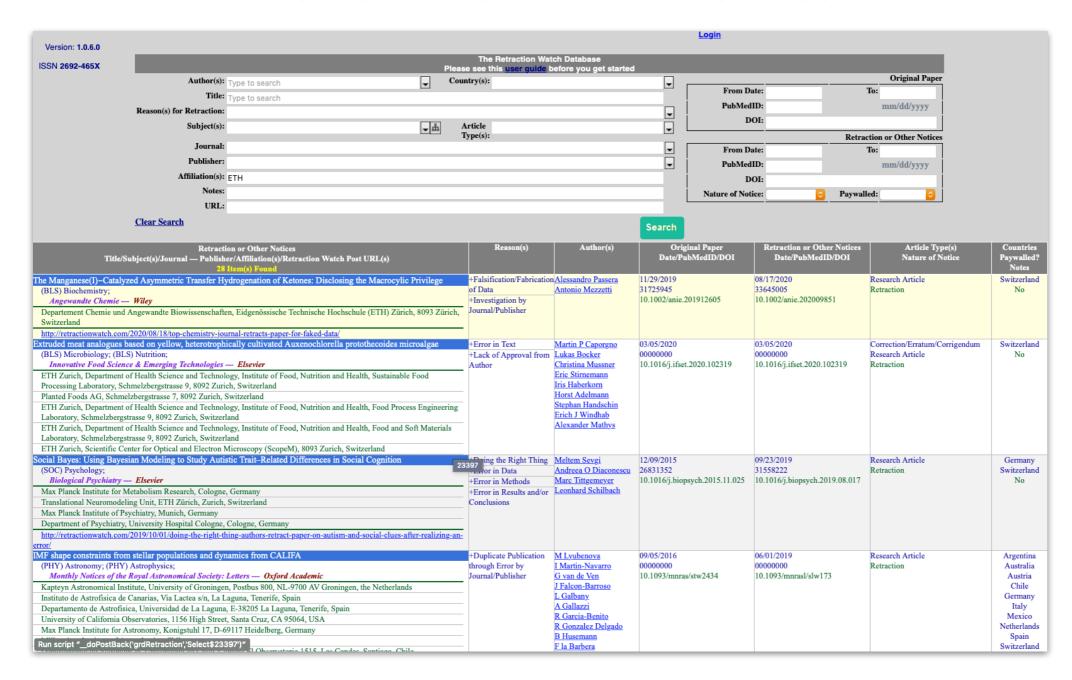
Retraction Watch

Retracted coronavirus (COVID-19) papers

Retracted

- 1. "5G Technology and induction of coronavirus in skin cells," published in *Biological Regulators & Homeostatic Agents* on July 16, 2020, withdrawn on July 24, 2020. Our coverage here.
- 2. "A data-mining based analysis of traditional Chinese medicine in diagnosing and treating COVID-19," published on June 24, 2021 in *The Anatomical Record*; unknown when retracted.
- 3. "A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell," published on October 17, 2020 in the *Journal of Nanoparticle Research*; retracted on August 16, 2021.
- 4. "A Retrospective Analysis and Comparison of Prisoners and Community-Based Patients with COVID-19 Requiring Intensive Care During the First Phase of the Pandemic in West Texas," published on August 29, 2021 in the Journal of Primary Care & Community Health; retracted on July 16, 2021
- "A review of convalescent plasma transfusion in COVID-19: Old wine reserved for special occasions," published in Lung India on September 16, 2020; retracted December 31, 2020.
- "A Study of Potential SARS-CoV-2 Antiviral Drugs and Preliminary Research of Their Molecular Mechanism, Based on Anti-SARS-CoV Drug
 Screening and Molecular Dynamics Simulation," published on December 1, 2020 in the Journal of Computational Biology; retracted on June 25, 2021.
- "Acute kidney injury and collapsing glomerulopathy associated with COVID-19 and APOL1 high risk genotype," Abstract 111 and Abstract 621, both published in the *Journal of Investigative Medicine*; both <u>retracted</u> on April 1, 2021.

The Retraction Watch Database



http://retractiondatabase.org/RetractionSearch.aspx?



Predatory journals are easy to please. They seem to accept papers with little regard for quality, at a fraction of the cost charged by mainstream open-access journals. These supposedly scholarly publishing entities are murky operations, making money by collecting fees while failing to deliver on their claims of being open access and failing to provide services such as peer review and archiving.

Source: Moher et al. (2017) Stop this waste of people, animals and money. Nature 549, 23–25.

2 – Gold Open Access – same publishing process as above. The difference is that when an article is accepted for publication, the author/s or funder/s pay an Article Processing Charge (APC). The final version of the published article is then free to read for everyone. The APC to publish Gold Open Access in *Nature* is £8890.00/\$12290.00/€10290.00.

nature

Beall's List - Potential, possible, or probable predatory scholarly open-access publishers

Frontiers' peer review process is flawed. It is stacked in favor of accepting as many papers as possible in order to generate more revenue for the company. Frontiers is included on my list, and I recommend against publishing in its journals, which are rather expensive to publish in anyway.

Is frontiers in microbiology a predatory Journal?

The Frontiers journals use open peer review, where the names of reviewers of accepted articles are made public. As of 2017, 24 of their journals had impact factors. ... Some journals, such as Frontiers in Human Neuroscience or Frontiers in Microbiology are considered megajournals on their own.





For Better Science

BY LEONID SCHNEIDER, ON RESEARCH INTEGRITY, BIOMEDICAL ETHICS AND ACADEMIC PUBLISHING

Beall-listed Frontiers empire strikes back



BY LEONID SCHNEIDER SEPTEMBER 14, 2016 COMMENTS 65



Willy

December 1, 2018

I think that both views have their own right to co-exist. I published articles in Frotniers and agree that their review process is sloppy. I also published articles with Wiley and Elsevier and experienced their review process as biased, unfair, and not constructive.

I reviewed for Frontiers and admit that my criticism of trash manuscripts was dismissed. However, it was possible to withdraw from the reviewing process so not to be stained by junk papers. I reviewed for Elsevier, Wiley, AAAS, NPG and others. Elsevier surely has the highest trash fraction among the manuscripts. However, they are happy to reject such manuscripts. Still, Elsevier is steering towards junk status as well as their reviewing process is flawed by nepotism. Plus, I consider Elsevier predatory as they bully university into a loan-shark style subscription model (<a href="https://www.the-scientist.com/news-

<u>opinion/universities-in-germany-and-sweden-lose-access-to-elsevier-journals-64522).</u>

I agree with CVAK in that Frontiers gives room to unconventional authors and ideas – and I applaud them for this. They also waive fees if one can't afford them as they have done for me.

The bottom line is that publishers are just the executive branch of science. They scientific system with its focus on quantity enables all sorts of spam, regardless of the publishers. Just remember the Mozart effect

(https://www.nature.com/articles/365611a0) or the memory of water (https://www.nature.com/articles/333816a0). Still, this is part of scientific discovery.

"[P]ublisher are just the executive branch of science."

Scientists publish their results to share them with the scientific community. It is common for peer-reviewed scientific journals to charge a fee for publishing a paper (Article Processing Charges). Depending on the journal, the cost of publication can be high. Recently, these fees have been questioned.

Interesting to read: Budzinski et al. (2020) Drivers of article processing charges in open access. Scientometrics 124, 2185–2206.

Example: From 2021, the publisher will charge €9,500 to make a paper open access (OA) in Nature.

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

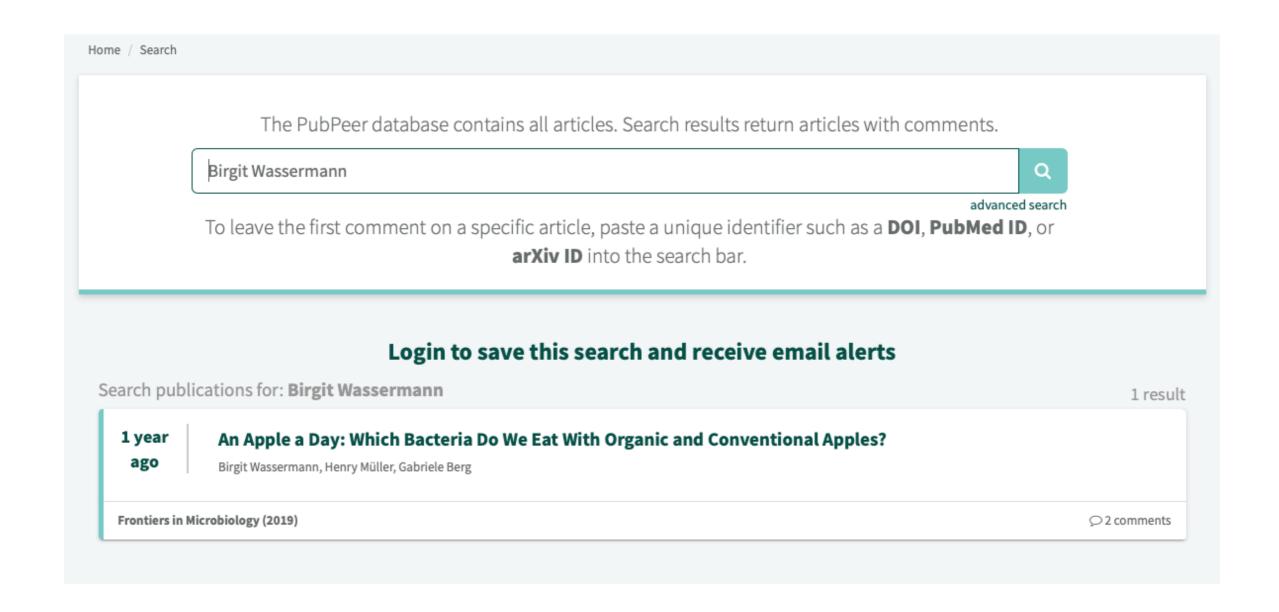
bioRxiv is an open access preprint repository for the biological sciences co-founded by John Inglis and Richard Sever in November 2013. It is hosted by the Cold Spring Harbor Laboratory. As preprints, papers hosted on bioRxiv are not peer-reviewed, but undergo basic screening and checked against plagiarism.

--Wikipedia

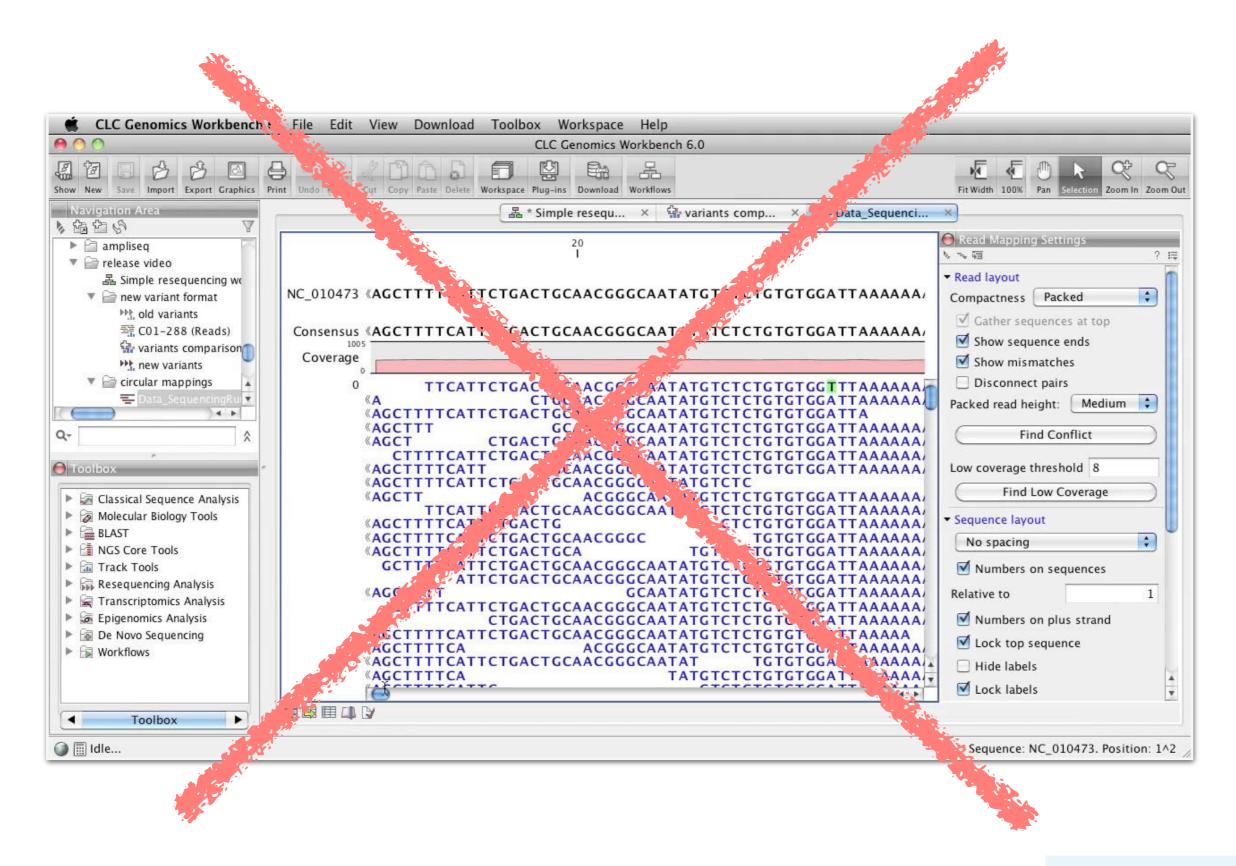


PubPeer is a website that allows users to discuss and review scientific research after publication, i.e. post-publication peer review.

--Wikipedia







Material and Methods (Manuscript)

In a first step, all paired-end raw reads were merged using FLASh (version 1.2.9, Magoc and Salzberg 2011) with minimum overlap of 5nt and maximal mismatch ration of 0.8.

Supplementary Data (Script or MD-Report)

```
## (a) Merging overlapping paired-end reads
# -v Version (1.2.9)
# -m minimum overlap (default 10bp)
# -x max mismatch ration (default 0.25)

flash -m 5 -x 0.8 random_1000_R1.fq random_1000_R2.fq -o
merged | tee flash.log
```



The Dryad Digital Repository is a curated resource that makes research data discoverable, freely reusable, and citable. Dryad provides a general-purpose home for a wide diversity of data types.

https://datadryad.org









WWW.PHDCOMICS.COM

```
error_correction.ph
error_correction_PE.ph
error_correction_PE_1.ph
error_correction_PE_new.ph
error_correction_SR_newer.ph
error_corection_SR_v190423.ph
error_correction_SR_newest.ph
```



Bioinformatics is not a messy buisness but requires organizational skills. There are many little (and for some people obvious) ways to get organized.

- File names
- Sufixes
- Special characters
- Temporary files
- Style

```
ReadME.txt
error_correction_SR_v190116.ph.archive
error_correction_SR_v190423.ph.archive
error_correction_SR_v190502.ph
error_correction_PE_v200317.ph.archive
error_correction_PE_v200421.ph
```

Question - Aim

Input file(s) - original and parsed

Program & Version (& Link)
Parameters (& References)

Output file(s) / Log-file(s)

Interpretation / Disscusion

Date: XX.YY.ZZ

Aim

Find differences between two nucleotide sequences.

Input

my file: Pram_sequence_A0021.fasta NCBI file: AY762091 (AY762091.fasta)

Pairwise alignment: LALIGN (Online Version 3.2.1)

http://www.ebi.ac.uk/Tools

Option: default



Results

```
Waterman-Eggert score: 682; 170.3 bits; E(1) < 1e-47
98.6% identity (98.6% similar) in 140 nt overlap (1-140:1-140)

10 20 30 40 50 60

A0021 ACACGTGCTACAATGGCCGTTACAGAGGGAAACCGCGAGGTGGAGCCAATCTCAG...

AY76091 ACACGTGCTACAATGGCCGTTACAGAGGGATTCGAAACCGCGAGGTGGAGCCAATCTCAG...
10 20 40 50 60
```

Discussion

My sequence (Pram sequence) aligns nicely with AY762091 from the NCBI database. There are, however, two nucleotide changes (red box).

Markdown Editor

```
### Merge Reads

In a first step, all paired-end raw reads were successfully merged using **FLASh** (version 1.2.11, Magoc and Salzberg 2011) with minimum overlap of 5nt and maximal mismatch ration of 0.8.

**``bash

## (a) Merging overlapping paired-end reads

# Version

flash -v | head -n 1

# Merge R1 and R2 reads

flash -m 5 -x 0.8 random 1000 R1.fg random 1000 R2.fg -o merged | tee flash.log

# Parameters:

# -m minimum overlap (default 10bp)

# -x max mismatch ration (default 0.25)

**``bash

grep "^@M" -c random_1000_R[12].fg

grep "^@M" -c random_1000_merged.fq

**```
```

HTML / PDF Report

Merge Reads

In a first step, all paired-end raw reads were successfully merged using **FLASh** (version 1.2.11, Magoc and Salzberg 2011) with minimum overlap of 5nt and maximal mismatch ration of 0.8.

```
## (a) Merging overlapping paired-end reads
# Version
flash -v | head -n 1
# Merge R1 and R2 reads
flash -m 5 -x 0.8 random_1000_R1.fq random_1000_R2.fq -o merged | tee flash.log
# Parameters:
# -m minimum overlap (default 10bp)
# -x max mismatch ration (default 0.25)
```

The merging rate was 87%.

```
grep "^@M" -c random_1000_R[12].fq
grep "^@M" -c random_1000_merged.fq
```

Headers

Headers are set using a hash before the title. The number of hashes before the title text will determine the depth of the header. Header depths are from 1-6

```
H1: # Header 1
H2: ## Header 2
H3: ### Header 3
H4: #### Header 4
H5: ##### Header 5
H6: ##### Header 6
```

Text Styling

```
Links: [Title](URL)
Bold: **Bold**
Italicize: *Italics*
Strike-through: ~~text~~
Highlight: ==text==
Paragraphs: Line space between paragraphs
Line break: Add two spaces to the end of the line
Lists: * an asterisk for every new list item.
Quotes: > Quote
Inline Code: alert('Hello World');
Horizontal Rule (HR): \------
```







Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.

A Quick Recap



Ask questions!

Not hesitate to ask for help if you have problems understanding or using a published method.

Reproducibility

Avoid applications with GUIs and use terminal command instead.



Provide Script / Write Reports

Precise description of the workflow including versions and parameters.



Share your findings!

Think carfully about where to publish your work and do it for the right reasons.

57



Zen