# Root cause of COVID-19? Biotechnology's dirty secret: Contamination. Bioinformatics evidence demonstrates that SARS-CoV-2 was created in a laboratory, unlikely to be a bioweapon but most likely a result of sloppy experiments

Vinu Arumugham
Apr 2020
vinucubeacc@gmail.com

## Abstract

Researchers at the Wuhan Institute of Virology (WIV) were using Human Immunodeficiency Virus (HIV-1) derived plasmids, cotransfecting them along with SARS-like virus derived spike gene plasmids, into human embryonic kidney (HEK) cells. This created pseudoviruses so that the experiment can be performed in a Biosafety Level 2 (BSL-2) laboratory instead of the BSL-4 required for SARS-like viruses. Their spike gene plasmid solution however, could of course be contaminated with SARS-like viruses due to inadequate purification, which infect the HEK cells. With HEK cells containing transfected HIV-1 genetic material, novel recombinant SARS-like viruses can be created with HIV-1 derived inserts. Such novel coronaviruses would have been exposed to laboratory workers and animals during handling, especially in a relaxed BSL-2 setting, thus resulting in SARS-CoV-2 and COVID-19.

Researchers identified the signature of 4 such HIV-1-like inserts in the SARS-CoV-2 spike protein. Two more inserts identified in the SARS-CoV-2 ORF1a protein are described here. All inserts have high homology to HIV-1 sequences.

Contamination is a fundamental problem that plagues drugs, vaccines, biologics, diagnostics and laboratory research materials. The biotechnology and pharmaceutical industries have done next to nothing about it. Egg, milk protein contaminated vaccines cause egg/milk allergy. Animal/plant/fungal protein contaminated vaccines cause the explosion of autoimmune disorders. Insulin injections are contaminated with yeast proteins which results in de novo autoimmune disorders. Rotavirus vaccines are contaminated with viable porcine circoviruses. Animal tissue and serum used in vaccine manufacturing are derived from animals infected by numerous diseases. And of course coronavirus infected animal material contaminating vaccines, cause IgE mediated sensitization against coronavirus proteins thus leading to COVID-19 severity. Products used in humans are supposed to be held to a "higher standard". So what level of contamination can we expect in diagnostics and materials used in laboratory research?

The WIV was experimenting with numerous SARS-like coronaviruses (SARS-L-CoV). The HEK were grown using fetal calf serum. So their experiments were contaminated with SARS-L-CoV and its genetic material, HIV-1 and its genetic material, any bovine viruses in the fetal calf serum, human viruses in HEK and their genetic material. When a cell is infected with multiple viruses or contaminated with genetic material from multiple viruses, recombination can occur creating novel viruses. With millions of cells contaminated as above, the Wuhan laboratory (and other such laboratories) were incubators for novel viruses including coronaviruses such as SARS-CoV-2. The laboratories in China were also notorious for leaking viruses. So it was only a matter of time before a virus with human-to-human transmission capability would leak. It did, in November 2019.

Biosafety is a misnomer. These are biohazard laboratories. These laboratories, around the world, must be immediately shut down before they wipe out humanity with the next disease.

## Introduction

A novel coronavirus, now named SARS-CoV-2 was identified in hospitalized patients in Wuhan, China, in December 2019. The disease caused by the virus, now named COVID-19, can range in severity from asymptomatic to acute respiratory distress syndrome (ARDS) and death.

Since SARS-like viruses are dangerous, they are difficult to study. One technique researchers use, is developing safer pseudoviruses that carry less risk of causing disease. Pseudoviruses are genetically modified organisms. Plasmids are circular DNA constructs that have instructions to make most of the essential parts of a virus and the capability to instruct a host cell to make copies of the pseudovirus. The SARS-like viruses use a spike protein to enter the host cell. Combining a spike gene plasmid and the HIV-1 derived plasmid, inside a cell (cotransfecting), one can create pseudoviruses that contain the spike protein just like a real SARS-like virus. Such a pseudovirus can infect cells just like a real virus but can do so only once. Plasmids can be derived from many viruses. The human immunodeficiency virus (HIV), is commonly used to create plasmids. Researchers at the Wuhan Institute of Virology (WIV) have been using this technique at least as early as the SARS outbreak in 2003.

WIV was using HIV-1 derived, pNL4.3.Luc.R−E−-Luc plasmids (1,2), cotransfecting them along with SARS-like virus derived spike gene plasmids, into human embryonic kidney (HEK) cells. This created pseudoviruses so that the experiment can be performed in a Biosafety level 2 (BSL-2) laboratory instead of the BSL-4 laboratory that would otherwise be required for handling SARS-like viruses (3). The spike gene plasmid solution is derived by purifying material obtained from a fecal swab of an infected bat. This spike gene plasmid solution however, could of course be contaminated with SARS-like viruses. These viruses can infect the HEK cells (4). With HEK cells containing transfected HIV-1 genetic material, novel recombinant SARS-like viruses can be created with HIV-1 derived inserts. Such novel coronaviruses would have been exposed to laboratory workers and laboratory animals during handling, especially in a relaxed BSL-2 setting.

## Discussion

### Contamination

Contamination is a fundamental problem that plagues drugs, vaccines, biologics, diagnostics and laboratory research materials. The biotechnology and pharmaceutical industries have done next to nothing about it. Egg, milk, peanut, soy, wheat, corn, fish, gelatin, sesame protein contaminated vaccines cause the development of food allergy (5–9). Animal/plant/fungal protein contaminated vaccines cause the explosion of autoimmune disorders (10,11). Aeroallergen contaminated vaccines cause the development of asthma and allergies (12). Insulin injections are contaminated with yeast proteins which results in de novo autoimmune disorders (13). Rotavirus vaccines are contaminated with viable porcine circoviruses (14). Animal tissue and serum used in vaccine manufacturing are derived from animals infected by numerous diseases. And of course, coronavirus infected animal material contaminating vaccines, cause IgE mediated sensitization directed against coronavirus proteins thus leading to COVID-19 severity (15). Biologics are contaminated with host cell proteins. These are usually Chinese Hamster Ovary (CHO) cell proteins. There is even a tool to analyze the immunogenecity danger of such contaminants, called CHOPPI (16). The Pandemrix vaccine was contaminated with H1N1 nucleoproteins that resulted in the narcolepsy disaster (17). The Arepanrix vaccine (like most influenza vaccines) contains numerous chicken proteins, 293 of which were identified (18). Carcinogenic N-Nitrosodimethylamine (NDMA) contaminates Ranitidine and Valsartan drugs. No one wants to talk about this dirty secret of the pharmaceutical and biotechnology industries.

If the drug stops working and affects revenue, as was the case with anti-drug antibodies (19) against contaminants in biologics, only then the manufacturer pays attention. Similarly, if the laboratory experiment produces the desired pseudoviruses, scientists are happy to ignore whatever else may be growing in their well plates.

Products used in humans are supposed to be held to a “higher standard”. So what level of contamination can we expect in diagnostics and materials used in laboratory research? Plasmid reagents can be contaminated with viruses that were used to derive them as well as other unrelated plasmids (3,20).

**SARS-like virus experiments using HIV-1 derived material**

The Wuhan Institute of Virology was experimenting with numerous SARS-like coronaviruses (SARS-L-CoV). They were using HIV-1 derived pseudoviruses, genetically engineered with the SARS-L-CoV spike genes. These experiments were being performed on human embryonic kidney (HEK) cells. The HEK were grown using fetal calf serum (1). So their experiments were contaminated with SARS-L-CoV and its genetic material, HIV-1 and its genetic material, any bovine viruses in the fetal calf serum, human viruses in HEK and their genetic material. When a cell is infected with multiple viruses or contaminated with genetic material from multiple viruses, recombination can occur creating novel viruses. With millions of cells contaminated as above in each experiment, the Wuhan laboratory (and other such laboratories) were incubators for novel viruses including coronaviruses such as SARS-CoV-2. The laboratories in China were also notorious for leaking viruses. So it was only a matter of time before a virus with human-to-human transmission capability would leak. It did, in November 2019.

Pradhan et al. identified the signature of 4 such HIV-1-like inserts in the SARS-CoV-2 spike protein (21). Two more such inserts identified in the SARS-CoV-2 ORF1a protein are described here. All inserts have high homology to HIV-1 sequences.

**Are these really inserts?**

Yes. The SARS-CoV-2 virus contains inserted sequences when compared to other SARS-like viruses. Sequences on either side of the inserts match. Pradhan et al. (21) identified four inserts in the spike protein with high homology to HIV-1 envelope gp120 and Gag proteins. The HIV-1 derived pNL4.3.Luc.R−E−-Luc plasmids used in WIV includes the HIV-1 envelope gp120, Gag and Pol protein encodings (2). Insert 4, the furin cleavage site insert (22), also has high homology (89%) to a bovine papillomavirus (BPV) E1 protein (BAF95810.1). BPV could be introduced as a contaminant of the fetal calf serum used as growth medium for HEK cells. Viruses can infect cells of other species. In nature, such infections may be less common due to the host’s immune system. In a laboratory setting with naked HEK cells, there is no protective immune barrier so infection by viruses from other hosts is more likely.

**More inserts**

The previously identified insert motifs are GTNGTKR, HKNNKS, RSYLTPGDSSSG and QTNSPRRA. Extending the analysis to other SARS-CoV-2 proteins, two inserts (inserts 5,6) were identified in the ORF1a protein : TVGQQDGSEDNQTTTIQTIV and QVEQKIA.

Insert 5:

QII57177.1 vs. Uniprot P0C6X7 (SARS-CoV-2 vs. SARS)

```
Query  961   PLEFGASAETVRVEEEEEEDWLDDTTEQS------------------EIEP--EPEPTP  999
             PLEFGA++  ++ EEE+EEDWLDD ++Q+                    E++P  E E TP
Sbjct  959   PLEFGATSAALQPEEEQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTP  1018
```

Insert 6: A motif, QVEQKIA, between the same two proteins above was also identified.

Pradhan et al. (21) performed protein sequence analysis of these insert motifs against all viral genomes and identified matches to HIV-1. Most common cause of recombination resulting in such inserts is co-infection of a cell with two viruses (23). Some have pointed out that the insert sequences are also commonly found in numerous other organisms. However, for the reason described, only the sequences in the viral genomes are relevant.

**Insert origin and recombination**

Think of a person making copies of two documents. If they are not careful, some pages of one document may end up in the other. This can happen in cells that are making copies of two viruses at the same time. Genetic material from one, gets inserted into another - known as recombination.

These inserts can occur only when genetic material from another organism is present in the same cell where the virus is replicating. In nature, this donor organism can be any bat virus (bat being the most common host for these coronaviruses), that can infect the same cells as the SARS-like virus. BLASTP searches show that the best matches to these insert sequences are to HIV-1 or BPV. These full insert sequences are not present in bat related viruses. These sequences have poor homology to bat viral proteins. So bat viruses could not have been the source of these inserts (Table 1).

In the laboratory, the inserts can occur when genetic material from another organism is introduced into a cell infected by the SARS-like virus. Both the introduction of genetic material and infection could be done on purpose or may be an unintended consequence of contamination.

In nature, for a bat virus to supply an insert to a SARS-like virus, the following conditions must be met:
The bat virus must contain the insert sequence in its genome.
The bat virus must be able to overcome barriers in the host and simultaneously infect the same cell as the SARS-like virus.
Six such sequential coinfections by the right viruses (see Table 1) have to occur to gather all the inserts identified in SARS-CoV-2.
These viruses must be able to infect that same species of bat and naturally occur in the same geographical area for such a sequence of coinfections to occur.

We don’t know the sequences in the specific HIV-1 derived plasmids used in the laboratory. Depending on the sequence, ALL these insert recombinations could have occurred during the course of as little as just one experiment, lasting 48 hours.(1) And we know WIV has been doing this for 17 years.

Table 1
SARS-CoV-2 insert match to laboratory viral material vs. wild bat viruses

| Insert Motif | Best match to lab viral material | Best match to a bat virus |
|---|---|---|
| GTNGTKR | envelope glycoprotein [Human immunodeficiency virus 1]<br><br>TNGTKR<br>TNGTKR | orf1ab polyprotein [Hipposideros bat coronavirus HKU10]<br><br>GTNGT<br>GTNGT |
| HKNNKS | envelope glycoprotein [Human immunodeficiency virus 1]<br><br>HKNNKS<br>HKNNKS | protein IV [Bat mastadenovirus G]<br><br>HKNN<br>HKNN |
| RSYLTPGDSSSG | envelope glycoprotein, partial [Human immunodeficiency virus 1]<br><br>RSYL---TPGDSSSG<br>RTYLFNETRGNSSSG | orf1ab polyprotein [Pipistrellus abramus bat coronavirus HKU5-related]<br><br>RSYLTP<br>RNYLTP |
| QTNSPRRA | E1 protein [Bovine papillomavirus type 9]<br><br>QTN-SPRRA<br>QTNNSPRRA<br><br>envelope glycoprotein [Human immunodeficiency virus 1]<br><br>QTNSPRRA<br>QTNS R A<br>QTNSSRXA | nucleoprotein [Bat coronavirus CDPHE15/USA/2006]<br><br>NSPRR<br>NSPRR |
| TVGQQDGSEDNQTTTIQTIV | envelope glycoprotein, partial [Human immunodeficiency virus 1]<br><br>GSEDNQTTTIQT<br>GSEDN T TI+T<br>GSEDNRTNTIET | fusion protein [Bat Paramyxovirus Eid_hel/GH-M74a/GHA/2009]<br><br>QDGSE-----DNQT<br>QDGSQTLMMIDNQT |
| QVEQKIA | gag protein, partial [Human immunodeficiency virus 1]<br><br>QVEQKI<br>QV+QKI<br>QVQQKI | non-structural polyprotein 1ab [Bat SARS-like coronavirus]<br><br>QVEQKIA<br>QVEQK+A<br>QVEQKVA |

As Pradhan et al. also point out, it is impossible for such inserts, all of them with such high levels of homology to HIV-1, to have naturally evolved in such a short time period (17 years since SARS). They wrote that it is a product of unconventional evolution. The parsimonious explanation for the inserts in SARS-CoV-2, is recombination with genetic material from HIV-1 derived plasmids used in the laboratory.

**The problem with the pedigree of RaTG13**

Zhou et al. (24) described a bat coronavirus named RaTG13 isolated in China's Yunnan province, which seems most closely related to SARS-CoV-2. This isolate from 2013 however, was only sequenced and uploaded to Genbank by the Wuhan Institute of Virology, after the COVID-19 outbreak. This raises many questions about RaTG13. Is it really a wild virus? Is it a cousin of SARS-CoV-2 from a laboratory? Many studies could have been mislead by RaTG13, on the origin of SARS-CoV-2 (23,25,26). So this virus must be ignored until we have independent isolation and sequencing to confirm if it really is a wild virus.

## Artificial selection at work

The WIV experiments with HIV-1 transfected HEK and SARS-like viruses were inadvertently supplying an opportunity for evolution via recombination with HIV-1 derived inserts and selecting the viruses for efficient human kidney cell infection.

## Could coinfection with HIV-1 and SARS in humans have created SARS-CoV-2?

Theoretically, if a human were infected simultaneously with HIV-1 and SARS viruses, recombination can occur. The inserts described above need not occur at once. They may have been separated by years. For recombination to occur in humans, SARS and HIV-1 have to infect the same cells. It is not known if HIV-1 and SARS can infect the same cells in humans. If it were possible and recombination occurs, SARS-like viruses with one or more HIV-1 like insert sequences would have caused outbreaks. They would have been isolated, sequenced and documented. No such viruses are known. It would be extremely unlikely that multiple inserts occurred at once in a human who also happens to live within miles of WIV, to explain the Wuhan outbreak. This would also run into difficulties explaining major changes in other parts of the SARS-CoV-2 virus in comparison to SARS.

## Conclusion

Biosafety is a misnomer. These are biohazard laboratories. This is obviously not just a problem in China. The US BSL-4 laboratory in Fort Detrick, Maryland was recently shut down by the Centers for Disease Control (CDC) due to safety concerns. These laboratories, around the world, must be immediately shut down before they wipe out humanity with the next disease they release.

Wild-life trade is an obvious risk factor for naturally evolving viruses that can jump to humans. Researchers often "gift" laboratory organisms, reagents, etc. - the lab-life trade - which can spread novel pathogens across the globe. So, along with wild-life trade, lab-life trade needs to be banned.

## Acknowledgment

The author thanks Dr. Meryl Nass, Maine, Dr. Loretta Bolgan, Italy and Dr. James Lyons-Weiler, for providing research material and for discussions.

**References**


1. Ren W, Qu X, Li W, Han Z, Yu M, Zhou P, et al. Difference in Receptor Usage between Severe Acute Respiratory Syndrome (SARS) Coronavirus and SARS-Like Coronavirus of Bat Origin. J Virol. American Society for Microbiology; 2008 Feb 15;82(4):1899–907.
2. Reagent Datasheet Detail: Catalog 3418 - HIV-1 NL4-3 ΔEnv Vpr Luciferase Reporter Vector (pNL4-3.Luc.R-E-) - NIH AIDS Reagent Program [Internet]. [cited 2020 Apr 19]. Available from: https://www.aidsreagent.org/reagentdetail.cfm?t=cloning_vectors&id=66
3. Li Q, Liu Q, Huang W, Li X, Wang Y. Current status on the development of pseudoviruses for enveloped viruses. Reviews in Medical Virology. John Wiley and Sons Ltd; 2018.
4. Yang Z-Y, Huang Y, Ganesh L, Leung K, Kong W-P, Schwartz O, et al. pH-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus Is Mediated by the Spike Glycoprotein and Enhanced by Dendritic Cell Transfer through DC-SIGN. J Virol. American Society for Microbiology; 2004 Jun 1;78(11):5642–50.
5. Yamane H. N. U. Serological examination of IgE- and IgG-specific antibodies to egg protein during influenza virus immunization. Epidemiol Infect. 1988;100(2):291–9.
6. Kattan JD, Cox AL, Nowak-Wegrzyn A, Gimenez G, Bardina L, Sampson HA, et al. Allergic reactions to diphtheria, tetanus, and acellular pertussis vaccines among children with milk allergy. J Allergy Clin Immunol. 2011;Conference(var.pagings):AB238.
7. Arumugham V. Evidence that Food Proteins in Vaccines Cause the Development of Food Allergies and Its Implications for Vaccine Policy. J Dev Drugs. OMICS International; 2015 Oct;04(04):1–3.
8. National Academies of Sciences and Medicine E. Finding a Path to Safety in Food Allergy: Assessment of the Global Burden, Causes, Prevention, Management, and Public Policy. Stallings VA, Oria MP, editors. Washington, DC: The National Academies Press; 2017.
9. Arumugham V, Trushin M V. Role of NMDA receptor autoimmunity induced by food protein containing vaccines, in the etiology of autism, type 1 diabetes, neuropsychiatric and neurodegenerative disorders [Internet]. International Journal of Pharmaceutical Research. 2019 [cited 2019 Apr 15]. p. 428–37. Available from: https://doi.org/10.5281/zenodo.1463600
10. Arumugham V. Vaccines and Biologics injury table based on mechanistic evidence – Feb 2020 Covering over 125 conditions [Internet]. 2020 [cited 2020 Apr 23]. Available from: https://doi.org/10.5281/zenodo.2582634
11. Hogenesch H, Azcona-Olivera J, Scott-Moncrieff C, Snyder PW, Glickman LT. Vaccine-induced autoimmunity in the dog. Adv Vet Med. 1999;41:733–47.

12. Arumugham V. Aeroallergen contamination of multi-dose and reconstituted vaccine vials cause the development of asthma, gastrointestinal diseases and proves vaccine makers and vaccine safety regulators are incompetent [Internet]. 2019 [cited 2019 Jan 22]. Available from: https://doi.org/10.5281/zenodo.2544037

13. Mankai A, Sakly W, Thabet Y, Achour A, Manoubi W, Ghedira I. Anti-Saccharomyces cerevisiae antibodies in patients with systemic lupus erythematosus. Rheumatol Int. Germany; 2013 Mar;33(3):665–9.

14. Vaccine Excipient & Media Summary [Internet]. 2018 [cited 2016 Jan 16]. Available from: http://www.cdc.gov/vaccines/pubs/pinkbook/downloads/appendices/B/excipient-table-2.pdf

15. Arumugham V. Immunological mechanisms explaining the role of IgE, mast cells, histamine, elevating ferritin, IL-6, D-dimer, VEGF levels in COVID-19 and dengue, potential treatments such as mast cell stabilizers, antihistamines, Vitamin C, hydroxychloroquine, ivermecti [Internet]. 2020 [cited 2020 Apr 24]. Available from: https://doi.org/10.5281/zenodo.3748303

16. Bailey-Kellogg C, Gutiérrez AH, Moise L, Terry F, Martin WD, De Groot AS. CHOPPI: a web tool for the analysis of immunogenicity risk from host cell proteins in CHO-based protein production. Biotechnol Bioeng. Wiley-Blackwell; 2014 Nov;111(11):2170–82.

17. Ahmed SS, Volkmuth W, Duca J, Corti L, Pallaoro M, Pezzicoli A, et al. Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. Sci Transl Med. 2015 Jul 1;7(294):294ra105–294ra105.

18. Jacob L, Leib R, Ollila HM, Bonvalet M, Adams CM, Mignot E. Comparison of Pandemrix and Arepanrix, two pH1N1 AS03-adjuvanted vaccines differentially associated with narcolepsy development. Brain Behav Immun. 2015 Jul;47:44–57.

19. Garcês S, Demengeot J. The Immunogenicity of Biologic Therapies. Curr Probl Dermatology. S. Karger AG; 2017;53:37–48.

20. Sun J, Tian Y, Du Y, Wang Z, Zhao G, Ma Y, et al. A cautionary tale of cross-contamination among plasmids from commercial suppliers. Biotechniques. Future Science Ltd; 2019 Nov 22;68(1):14–21.

21. Pradhan P, Pandey AK, Mishra A, Gupta P, Tripathi PK, Menon MB, et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag [Internet]. bioRxiv. Cold Spring Harbor Laboratory; 2020 [cited 2020 Apr 23]. p. 2020.01.30.927871. Available from: https://www.biorxiv.org/content/10.1101/2020.01.30.927871v1.full

22. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. Elsevier B.V.; 2020 Apr 1;176:104742.

23. Xiao C, Li X, Liu S, Sang Y, Gao SJ, Gao F. HIV-1 did not contribute to the 2019-nCoV genome. Emerging Microbes and Infections. Taylor and Francis Ltd.; 2020. p. 378–81.

24. Zhou P, Yang X Lou, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. Nature Research; 2020 Mar 12;579(7798):270–3.

25. Paraskevis D, Kostaki EG, Magiorkinis G, Panayiotakopoulos G, Sourvinos G, Tsiodras S. Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. Infect Genet Evol. Elsevier B.V.; 2020 Apr 1;79.

26. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med. Springer Science and Business Media LLC; 2020 Mar 17;26(4):450–2.