

CONFIDENTIAL

Date: July 24, 2019

TO: Dr. Lisa Cassis
Vice President for Research

Dr. David Blackwell
Provost

FROM: Sanda I. Despa, PhD, Department of Pharmacology and Nutritional Sciences (College of Medicine); Sylvie Garneau-Tsodikova, PhD, Department of Pharmaceutical Sciences (College of Pharmacy); and Sidney W. Whiteheart, PhD, FAHA, Department of Molecular and Cellular Biochemistry (College of Medicine)

RE: Allegation of Research Misconduct involving Drs. Xianglin Shi, Zhuo Zhang, and Donghern Kim (collectively, the “Respondents”)

Table of Contents

Section	Heading	page
A	INVESTIGATION COMMITTEE MEMBERS	2
B	CONSULTANTS	2
C	FUNDING SOURCE	2
D	FORMAL ALLEGATIONS	2
E	INITIATING EVENTS	2
F	CHARGE TO COMMITTEE	3
G	DEFINITION OF RESEARCH MISCONDUCT	3
H	OVERALL PROCESS	3
I	LIST OF RESEARCH RECORDS AND EVIDENCE REVIEWED	4
J	PROCEDURE AND DISCUSSION	6
K	LIST OF THE APPENDICES ASSOCIATED WITH THIS REPORT	10
L	SUMMARY STATEMENT OF FINDINGS	14
M	DETAILED ANALYSIS OF FINDINGS IN THE 10 SIGNIFICANT DEPARTURE CATEGORIES IDENTIFIED	17
M.1	Specific examples for “no original data provided”	18
M.2	Inappropriate loading controls	47
M.3	Inappropriate modification of original data	62
M.4	Inappropriate scale bars on images	75
M.5	Data provided did not match the published figure	82
M.6	Incomplete metadata provided with figure components	88
M.7	Black images containing no pixel data	101
M.8	Data fabrication and falsification	109
M.9	Falsified/fabricated data provided to the committee	117
M.10	Retracted manuscripts	119
N	Respondents’ responses to report	149
O	Committee remarks on Respondents’ comments	149

P	Recommendations	154
---	-----------------	-----

A. INVESTIGATION COMMITTEE MEMBERS

1. Sanda I. Despa, PhD, Department of Pharmacology and Nutritional Sciences
(College of Medicine)
2. Sylvie Garneau-Tsodikova, PhD, Department of Pharmaceutical Sciences
(College of Pharmacy)
3. Sidney W. Whiteheart, PhD, FAHA, Department of Molecular and Cellular Biochemistry
(College of Medicine)

B. CONSULTANTS

The committee did not use scientific consultants. However, the committee did consult with the University of Kentucky’s Helene Lake-Bullock, PhD, JD (Office of Research Integrity), David L. Kinsella, JD, and Theresa B. Crocker, PhD (each with the Office of Legal Counsel) for advice on investigative procedures and document generation.

C. FUNDING SOURCE

The original subject of this allegation is the NIH grant application, 3R01ES025515-03S1, titled “Oxidative stress, Cr(VI) carcinogenesis, and prevention” submitted to the National Institute of Environmental Health Sciences (NIEHS). This supplement was not funded. Dr. Shi was listed as the Principal Investigator, Dr. Zhang as the MPI, and Dr. Kim contributed data and information for the grant application.

D. FORMAL ALLEGATIONS

The University received an allegation that data were falsified in a supplemental grant application entitled “Oxidative stress, Cr(VI) carcinogenesis, and prevention” submitted to NIH/NIEHS. Specifically, the allegation was that figure panels purported to represent images captured by fluorescent microscopy were instead black squares which did not contain any data. This supplement has not been funded.

E. INITIATING EVENTS

1. On May 10, 2018, Complainant sent an email to Helene Lake-Bullock, questioning the data in an NIH grant application, 3R01ES025515-03S1, for which Dr. Shi was the University’s Principal Investigator.
2. An inquiry was then conducted and concluded with a recommendation to proceed to an investigation, and the inquiry committee’s report, including without limitation the Respondents’ response, was provided to the University of Kentucky Vice President for Research, Dr. Lisa Cassis.
3. On September 14, 2018, Dr. Lisa Cassis sent all three Respondents a letter of notification regarding an investigation of research misconduct.

4. Conflict of interest statements from potential committee members were collected.
5. On September 14, 2018, Dr. Cassis appointed Drs. Sidney Whiteheart, Sylvie Garneau-Tsodikova, and Vivek Rangnekar to investigate allegations of research misconduct against the Respondents. The Respondents objected to the inclusion of Dr. Rangnekar on the committee because he served on graduate committees of the Respondents' students, so on September 19, 2018, Dr. Sanda Despa was appointed in his place.

F. CHARGE TO COMMITTEE

The Investigation Committee was presented with the allegation that there was sufficient evidence and unresolved issues to warrant an investigation into whether the allegations that Figs. 6 and 10 of grant application 3R01ES025515-03S1 contained falsified data. The committee was directed to develop a factual record exploring the allegations in detail based upon the available evidence including the testimony of the Respondents and any key witnesses and exploring the evidence in depth, leading to a finding of whether research misconduct had been committed and if so by whom and to what extent.

G. DEFINITION OF RESEARCH MISCONDUCT

Research misconduct is defined as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results. Research misconduct does not include honest error or differences of opinion (Appendices 001 and 002).

H. OVERALL PROCESS

1. On September 17, 2018, the first organization meeting was held. Helene-Lake Bullock and David Kinsella provided Drs. Whiteheart, Garneau-Tsodikova, and Rangnekar with an overall outline of the investigation process and answered questions. Shortly thereafter, Dr. Rangnekar was removed from the committee after an objection by the Respondents, and Dr. Despa was appointed in his place. The Inquiry Report and additional materials from the preliminary inquiry were provided. Throughout the investigation, information pertinent to the investigation was collected. The only sources used were the Respondents and the data collected from them, others that were interviewed and the University's Offices of Research Integrity and Sponsored Projects Administration, and there was no analysis done outside the Investigation Committee. The members of the Investigation Committee also met on a number of occasions during this time to discuss aspects of the investigation process. *The time-line of events that comprise this investigation can be found in Section J.*
2. The Investigation Committee conducted in-person interviews with the following people (listed in chronological order):
Donghern Kim, Respondent, January 10, 2019 and January 31, 2019
Robert DiPaola, January 10, 2019 and January 24, 2019
Andrew Hitron, January 10, 2019
Xianglin Shi, Respondent, January 17, 2019

Zhuo Zhang, Respondent, January 17, 2019
Lei Wang, January 24, 2019

I. LIST OF RESEARCH RECORDS AND EVIDENCE REVIEWED

Based on their CVs and the grants provided to the committee by the Office of Sponsored Projects and Administration (OSPA), Drs. Shi, Zhang, and Kim published approximately 60 publications (Appendices 003 (Shi publications), 004 (Zhang publications, and 005 (Kim publications)) and submitted 7 R level NIH grant proposals during the period (2012-2018) that the committee investigated. The committee performed a general overview of these documents and identified 19 items (7 grants and 12 publications) to examine in more detail based on identified potential irregularities. The investigative committee requested that the Respondents supply supporting data and documents for these 19 items. Nine distinct classes of irregularities were noted in the material provided. A tabulation of the committee's findings is below. A specific narrative defining the irregularities and describing each incident follows in the committee's report.

In Grants&Publications.PDF:

Shi listed: (17 grants: 9 R01s, 2 R21s, 3 R13s, 3 P30s). Amongst those the committee received from OSPA: R01s #1 (our #3), #4 (our #2), #5 (our #1), #6 (our #6), and #7 (our #4) as well as P30s #2 and #3. The committee did not receive the other grants listed by Dr. Shi. The committee only examined the R01s that were provided to it by OSPA.

Zhang listed: (7 R01s, 1 R21, 1 R03). Amongst these, #3 (our #5) was additionally provided.

LIST OF 7 GRANTS (G1-G7) AND 12 MANUSCRIPTS (M1-M12) FOR WHICH THE COMMITTEE REQUESTED DATA AS WELL AS GENERAL DOCUMENTS (hard drives or other USB devices, lab notebooks, notepads, any hard copies or electronic documents where information related to experiments is recorded) REQUESTED:

Grants:

- G1. Grant Shi_3210000529 = 3R01ES025515-03S1 (02/01/2017-01/30/2020) (Appendix 006, pages 11-133): Figures 6 and 10
- G2. Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Figures 3, 9, 10, and 13
- G3. Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Figures 3, 6, 7, 8, and 9
- G4. Grant Zhang_3048111797 = 1R01ES021771-01A1 (08/01/2014-04/30/2019) (Appendix 009): Figure 3
- G5. Grant Zhang_3200001472 = 5R01ES028321-02 (09/01/2017-07/31/2022) (Appendix 010): Figures 10, 11, and 15
- G6. Grant Zhang_3200001638 = 1R01ES028984-01 (12/15/2017-11/30/2022) (Appendix 011): Figures 3, 8, and 11
- G7. Grant Zhang_3200001897 = 1R01CA228236-01A1 (01/01/2018-05/31-2023) (Appendix 012): Figures 4, 9, 13, 16, and 17

Manuscripts:

- M1. Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): all figures in manuscript.
- M2. Son YO, Pratheeshkumar P, Wang Y, Kim D, Zhang Z, and Shi X. (2017). Protection from Cr(VI)-induced malignant cell transformation and tumorigenesis of Cr(VI)-transformed cells by luteolin through Nrf2 signaling. *Toxicology and Applied Pharmacology*, 331, 24-32 (Appendix 014): all figures in manuscript.
- M3. Pratheeshkumar P, Son YO, Divya SP, Wang L, Turcios L, Roy RV, Hitron JA, Kim D, Dai J, Asha P, Zhang Z, and Shi X. (2017). Quercetin inhibits Cr(VI)-induced malignant cell transformation by targeting miR-21-PDCD4 signaling pathway. *Oncotarget*, 8, 52118-52131 (Appendix 015): all figures in manuscript.
- M4. Gao, N., Cheng, S., Budhraj, A., Liu, E.H., Chen, J., Chen, D., Yang, Z., Luo, J., Shi, X., and Zhang, Z. 3,3'-Diindolylmethane exhibits antileukemic activity in vitro and in vivo through a Akt-dependent process. *PLoS One*. 7, e31783, 2012 (Appendix 016): all figures in manuscript.
- M5. Wang, L., Son, Y.O., Ding, S., Wang, X., Hitron, J.A., Budhraj, A., Lee, J.C., Lin, Q., Poyil, P., Zhang, Z., Luo, J., and Shi, X. Ethanol enhances tumor angiogenesis in vitro induced by low-dose arsenic in colon cancer cells through hypoxia-inducible factor 1 alpha pathway. *Toxicological Sciences* 130, 269-280, 2012 (Appendix 017): Figure 2C.
- M6. Yang, Y., Wang, H., Wang, S., Xu, M., Liao, M., Frank, J.A., Adhikari, S., Bower, K.A., Shi, X., Ma, C., and Luo, J. GSK3 β signaling is involved in ultraviolet B-induced activation of autophagy in epidermal cells. *International Journal of Oncology* 41, 1782-1788, 2012 (Appendix 018): Figures 3C and 6B.
- M7. Yin, Y., Li, W., Son, Y.O., Sun, L., Kim, D., Wang, X., Yao, H., Wang, L., Pratheeshkumar, P., Hitron, A., Luo, J., Gao, N., and Shi, X., and Zhang, Z. Quercitrin protects skin from UVB-induced oxidative damage. *Toxicology and Applied Pharmacology* 269, 89-99, 2013 (Appendix 019): Figures 1E, 1F, 5A, 5B, and 7C.
- M8. Wang, L., Kung, L., Hitron, J.A., Son, Y.O., Wang, X., Budhraj, A., Lee, J.C., Pratheeshkumar, P., Chen, G., Zhang, Z., Luo, J., and Shi, X. Apigenin suppresses migration and invasion of transformed cells through down-regulation of C-X-C chemokine receptor 4 expression. *Toxicology and Applied Pharmacology* 272, 108-116, 2013 (Appendix 020): Figure 4C.
- M9. Kim, D., Dai, J., Park, Y.H., Yenwong F., L., Wang, L., Pratheeshkumar, P., Son, Y.O., Kondo, K., Xu, M., Luo, J., Shi, X., and Zhang, Z. Activation of EGFR/p38/HIF-1 α is pivotal for angiogenesis and tumorigenesis of malignantly transformed cells induced by hexavalent chromium. *Journal of Biological Chemistry* 291, 16271-16281, 2016 (Appendix 021): Figures 2Hb, 2Ib, and 6Bb.
- M10. Ren, Z., Yang, F., Wang, X., Wang, Y, Xu, M, Frank, J.A., Ke, Z.J., Zhang, Z., Shi, X., Luo, J. Chronic plus binge ethanol exposure causes more severe pancreatic injury and inflammation. *Toxicology and Applied Pharmacology* 308, 11019, 2016 (Appendix 022): Figure 3C.
- M11. Roy, R.V., Pratheeshkumar, P., Son, Y.O., Wang, L., Hitron, J.A., Divya, S.P., Zhang, Z., Shi, X., Different roles of ROS and Nrf2 in Cr(VI)-induced inflammatory responses in normal and Cr(VI)-transformed cells. *Toxicology and Applied Pharmacology* 307, 81-90, 2016 (Appendix 023): all figures in manuscript.

M12. Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROSdependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Figures 2D and 3E.

J. PROCEDURE AND DISCUSSION

1. On September 4, 2018, Dr. Lisa Cassis informed the Federal Office of Research Integrity that the University of Kentucky (UK) would move forward with an investigation against the Respondents (Appendix 025).
2. On September 14, 2018, the Respondents were informed that an Investigative Committee would be appointed to look into allegations of research misconduct (Appendices 026 (for Shi), 027 (for Zhang), and 028 (for Kim)).
3. On September 17, 2018, the Investigation Committee (Drs. Whiteheart, Garneau-Tsodikova, and Rangnekar, who was shortly replaced with Dr. Despa) met with Helene Lake-Bullock, PhD, JD and David Kinsella, JD for an orientation meeting. The group went over the review process and discussed University of Kentucky Administrative Regulations (Appendix 002) and Federal (DHHS) policies regarding research misconduct (Appendix 001). The Inquiry Committee report (Appendix 006) was also reviewed.
4. On September 17, 2018, the committee requested CVs from each of the Respondents (Appendices 029 (for CVs request) and 030 (Shi CV), 031 (Zhang CV), and 032 (Kim CV)).
5. On September 18, 2018, a request was received from Dr. Shi to remove Dr. Rangnekar from the committee because of a perceived conflict of interest. Dr. Despa was appointed in his place.
6. On October 15, 2018, the committee requested that Dr. Kim provide it with a demonstration of his use of the lab microscope (Appendix 033).
7. On October 18, 2018, the committee presented its first request for data to the Respondents, in which it asked for raw data and images for figure in 6 grants and 3 manuscripts. The committee requested that this data be produced by October 22, 2018 (Appendix 034).
8. On October 19, 2018, the committee observed Dr. Kim using the lab microscope. The demonstration was recorded (Appendix 035, DVD recording).
9. On October 22, 2018, Respondents requested and received an extension to the October 22, 2018 deadline for the first production of data. A new deadline of October 23, 2018 was granted. Respondents also objected to the scope of the request as being outside the original allegations of research misconduct (Appendix 036).
10. On October 24, 2018, the committee granted the Respondents a further extension to October 26, 2018 to respond to the first request for data (Appendix 037).

11. On October 24, 2018, the committee presented its second request for data, including general lab materials, to the Respondents. The committee requested that this data be produced by November 2, 2018 (Appendix 037).
12. On October 26, 2018, the Respondents provided the responses to the first request for data from the committee, as well as lab notebooks and hard copy documents that the committee requested on October 24, 2018 (Appendix 038).
13. On October 30, 2018, the Respondents provided some updates in response to committee requests 2 and 8 (Appendix 039).
14. On November 2, 2018, the Respondents provided data in response to the committee's second request for data (Appendix 040).
15. On November 9, 2018, the committee requested that the Respondents provide data on which of their grants were funded by the NIH and which of their manuscripts contained research results from grants funded by the NIH (Appendix 041). The Respondents provided the data in response to this request on November 16, 2018 (Appendices 042 and 043).
16. On December 3, 2018, Dr. Lisa Cassis requested a 120-day extension from the Federal Office of Research Integrity to complete the investigation of the Respondents, for a final due date of the report of April 14, 2019 (Appendix 044).
17. On December 10, 2018, the committee made a third request for complete data, including hard copies, protocols, and computer file paths for four figures, two from grants and two from manuscripts. The committee requested that this data be produced by December 14, 2018 (Appendix 045).
18. On December 14, 2018, the Respondents provided data in response to the committee's third request (Appendix 046).
19. On December 17, 2018, the Respondents provided further data in response to the first committee request for data on October 18, 2018 for Figs. 3 and 6 of grant Zhang_3048111797 (Appendix 047).
20. On January 10, 2019, Dr. DiPaola was interviewed (Transcript, Appendix 048). Subjects of the interview included (i) the nature and details of his collaboration with the Respondents, and (ii) his role in evaluating the data and generating and handling manuscript M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).
21. On January 10, 2019, Drs. Kim (Transcript, Appendix 049 and Exhibits, Appendix 050) and Hitron were interviewed (Transcript, Appendix 051). Subjects of the interviews included (i) training in the laboratory, (ii) procedures for recording and saving experimental details, (iii) their interactions with Drs. Shi, Zhang, and DiPaola, (iv) general laboratory practices for

writing manuscripts and grants and generating the figures that go in these documents, (v) defining what is a valid load control, and (vi) laboratory practices for resolving conflicts in data.

22. On January 17, 2019, Drs. Shi (Transcript, Appendix 052 and Exhibits, Appendix 053) and Zhang (Transcript, Appendix 054 and Exhibits, Appendix 053)¹ were interviewed. Subjects of the interviews included (i) description of laboratory organization and training of laboratory members, (ii) practices for producing manuscripts including figures, text, experimental protocols, and list of authors, (iii) practices for writing grants including figures, text, and experimental details, (iv) how original data are confirmed, (v) the nature and details of the Respondents' collaboration with Dr. DiPaola and their roles in evaluating the data and generating and handling manuscript M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013), (vi) laboratory practices for resolving conflicts in data, (vii) sequence of events that led to the retraction of three manuscripts from the *Journal of Biological Chemistry* below (R1. Son, Y.O., Pratheeshkumar, P., Divya, S.P., Zhuo Zhang, Z., and Shi, X. Nuclear factor erythroid 2-related factor 2 enhances carcinogenesis by suppressing apoptosis and promoting autophagy in nickel-transformed cells. *Journal of Biological Chemistry*, 292, 8315-8330, 2017 (Appendix 055); R2. Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Divya, S.P., Xu, M., Luo, J., Chen, G., Zhang, Z. and Shi, X. Antioncogenic and oncogenic properties of Nrf2 in arsenic-induced carcinogenesis. *Journal of Biological Chemistry*, 290, 27090-27100, 2015 (Appendix 056); R3. Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Zhang, Z., and Shi, X. Nrf2/p62 signaling in apoptosis resistance and its role in cadmium-induced carcinogenesis. *Journal of Biological Chemistry*, 293, 15455, 2014 (Appendix 057)), (viii) inability to produce original raw data, and (ix) whether they were aware of UK's and NIH's regulations on retaining original research data.
23. On January 24, 2019, Dr. Wang was interviewed (Transcript, Appendix 058 and Exhibits, Appendix 059). Subjects of the interview included (i) procedures for recording and saving experimental details, (ii) the nature and details of his collaboration with Dr. DiPaola, and (iii) his role in evaluating the data and generating manuscript M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).
24. On January 24, 2019, Dr. DiPaola was interviewed for the second time (Transcript, Appendix 060 and Exhibits, Appendix 061). Subjects of the interview included (i) establishing who was in charge of financial support and supervision of Dr. Wang, and (ii) financing of other people in other laboratories.
25. On January 29, 2019, Dr. Shi forwarded additional materials relative to Fig. 10 in grant application G1 = Shi_3210000529 = 3R01ES025515-03S1 from the original allegation (Appendix 062).

¹ The exhibits for both the Shi and Zhang interviews are the same.

26. On January 31, 2019, Dr. Kim was interviewed for the second time (Transcript, Appendix 063 and Exhibits, Appendix 064). Subjects of the interviews included (i) his role in evaluating the data and generating manuscript M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013), (ii) the nature of his interactions with Dr. DiPaola, and (iii) detailed discussion of Fig. 4C from manuscript M1.
27. On February 12, 2019, Dr. Shi forwarded additional materials relative to Figs. 6 and 10 in grant application G1 = Shi_3210000529 = 3R01ES025515-03S1 from the original allegation (Appendix 065).
28. On February 19, 2019, the Respondents were notified regarding the expansion of the original allegations of research misconduct (Appendices 066 (for Shi), 067 (for Zhang), and 068 (for Kim)).
29. On February 21, 2019, Dr. Kim provided additional materials regarding Fig. 4C (Appendix 069) in manuscript M1 = Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).
30. On February 22, 2019, the committee requested that the Respondents send them no further documentation or data relative to the investigation until after they received the committee's report (Appendix 070).
31. On March 25, 2019, Dr. Lisa Cassis requested an extension of 60 days to complete UK's investigation of the Respondents, for a final due date for the report of June 14, 2019 (Appendix 071).
32. On April 8, 2019, Respondents were notified of a further expansion of the investigation (Appendices 072 (for Shi), 073 (for Zhang), and 074 (for Kim)).
33. On April 12, 2019, Dr. Shi sent information to Dr. Lisa Cassis in partial response to the April 8, 2019 notification. (Appendix 075).
34. On April 15, 2019, Dr. Shi sent a further response to the April 8, 2019 notification to Dr. Lisa Cassis. (Appendix 076).
35. On April 15, 2019, counsel for Drs. Shi and Zhang, Bernard Pafunda, JD, shared a letter (from Yan Chen of the Shanghai Institute of Nutrition and Health to Dr. Shi of April 11, 2019) with David L. Kinsella as an additional response to the April 8, 2019 notification of Dr. Lisa Cassis (Appendix 077).
36. On April 16, 2019, the committee requested the raw data used to produce the figures in the three manuscripts that Drs. Shi and Zhang had published in the *Journal of Biological Chemistry*, which were later (on September 7, 2018) withdrawn by Drs. Shi and Zhang. The committee requested that this data be produced by April 18, 2019 (Appendix 078).

37. On April 18, 2019, Dr. Shi informed the committee that the Respondents would not be able to meet the deadline for producing the data requested on April 16, 2019 (Appendix 079).
38. On April 24, 2019, the committee extended the deadline to provide the data in response to their April 16, 2019 request to April 29, 2019 (Appendix 080).
39. On April 26, 2019, the committee requested that the Respondents provide information regarding any non-US funding or affiliation associated with their research or publications. The committee requested that this information be produced by May 6, 2019 (Appendix 081).
40. On April 29, 2019, Drs. Shi and Zhang produced some of the data requested by the committee on April 16, 2019 (Appendix 082).
41. On May 3, 2019, the Respondents produced information regarding any non-US funding or affiliations associated with their research or publications (Appendices 083 (email), 084 (for Shi), 085 (for Zhang), and 086 (for Kim)).

K. LIST OF THE APPENDICES ASSOCIATED WITH THIS REPORT

Number	Title
001	42 C.F.R. . § 93.103
002	University of Kentucky Administrative Regulation 7:1, Research Misconduct
003	List of Dr. Shi publications
004	List of Dr. Zhang publications
005	List of Dr. Kim publications
006	Inquiry report
007	Grant, Shi 3048112536
008	Grant, Shi 3200001792
009	Grant, Zhang 3048111797
010	Grant, Zhang 3200001472
011	Grant, Zhang 3200001638
012	Grant, Zhang 3200001897
013	Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. <i>The Prostate</i> , 78, 390-400
014	Son YO, Pratheeshkumar P, Wang Y, Kim D, Zhang Z, and Shi X. (2017). Protection from Cr(VI)-induced malignant cell transformation and tumorigenesis of Cr(VI)-transformed cells by luteolin through Nrf2 signaling. <i>Toxicology and Applied Pharmacology</i> , 331, 24-32
015	Pratheeshkumar P, Son YO, Divya SP, Wang L, Turcios L, Roy RV, Hitron JA, Kim D, Dai J, Asha P, Zhang Z, and Shi X. (2017). Quercetin inhibits Cr(VI)-induced malignant cell transformation by targeting miR-21-PDCD4 signaling pathway. <i>Oncotarget</i> , 8, 52118-52131
016	Gao, N., Cheng, S., Budhraj, A., Liu, E.H., Chen, J., Chen, D., Yang, Z., Luo, J., Shi, X., and Zhang, Z. 3,3'-Diindolylmethane exhibits antileukemic activity in vitro and in vivo through a Akt-dependent process. <i>PLoS One</i> . 7, e31783, 2012
017	Wang, L., Son, Y.O., Ding, S., Wang, X., Hitron, J.A., Budhraj, A., Lee, J.C., Lin, Q., Poyil, P., Zhang, Z., Luo, J., and Shi, X. Ethanol enhances tumor angiogenesis in vitro induced by low-dose

	arsenic in colon cancer cells through hypoxia-inducible factor 1 alpha pathway. <i>Toxicological Sciences</i> 130, 269-280, 2012
018	Yang, Y., Wang, H., Wang, S., Xu, M., Liao, M., Frank, J.A., Adhikari, S., Bower, K.A., Shi, X., Ma, C., and Luo, J. GSK3 β signaling is involved in ultraviolet B-induced activation of autophagy in epidermal cells. <i>International Journal of Oncology</i> 41, 1782-1788, 2012
019	Yin, Y., Li, W., Son, Y.O., Sun, L., Kim, D., Wang, X., Yao, H., Wang, L., Pratheeshkumar, P., Hitron, A., Luo, J., Gao, N., and Shi, X., and Zhang, Z. Quercitrin protects skin from UVB-induced oxidative damage. <i>Toxicology and Applied Pharmacology</i> 269, 89-99, 2013
020	Wang, L., Kung, L., Hiltron, J.A., Son, Y.O., Wang, X., Budhraj, A., Lee, J.C., Pratheeshkumar, P., Chen, G., Zhang, Z., Luo, J., and Shi, X. Apigenin suppresses migration and invasion of transformed cells through down-regulation of C-X-C chemokine receptor 4 expression. <i>Toxicology and Applied Pharmacology</i> 272, 108-116, 2013
021	Kim, D., Dai, J., Park, Y.H., Yenwong F., L., Wang, L., Pratheeshkumar, P., Son, Y.O., Kondo, K., Xu, M., Luo, J., Shi, X., and Zhang, Z. Activation of EGFR/p38/HIF-1 α is pivotal for angiogenesis and tumorigenesis of malignantly transformed cells induced by hexavalent chromium. <i>Journal of Biological Chemistry</i> 291, 16271-16281, 2016
022	Ren, Z., Yang, F., Wang, X., Wang, Y., Xu, M., Frank, J.A., Ke, Z.J., Zhang, Z., Shi, X., Luo, J. Chronic plus binge ethanol exposure causes more severe pancreatic injury and inflammation. <i>Toxicology and Applied Pharmacology</i> 308, 11019, 2016
023	Roy, R.V., Pratheeshkumar, P., Son, Y.O., Wang, L., Hitron, J.A., Divya, S.P., Zhang, Z., Shi, X., Different roles of ROS and Nrf2 in Cr(VI)-induced inflammatory responses in normal and Cr(VI)-transformed cells. <i>Toxicology and Applied Pharmacology</i> 307, 81-90, 2016
024	Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROSdependent activation of STAT3-miR-21-PDCD4 mechanism. <i>Scientific Reports</i> 6, 37227, 2016
025	September 4, 2018 letter from Lisa Cassis, UK VP for Research to Ranjini Ambalavanar, Acting Division Director, HHS/ORI
026	September 14, 2018 letter from Lisa Cassis to Xianglin Shi
027	September 14, 2018 letter from Lisa Cassis to Zhou Zhang
028	September 14, 2018 letter from Lisa Cassis to Donghern Kim
029	September 17, 2018 request from committee (CVs of Respondents)
030	Xianglin Shi CV
031	Zhou Zhang CV
032	Donghern Kim CV
033	October 15, 2018 request from committee (Demonstration of laboratory microscope)
034	October 18, 2018 request from committee (First data request)
035	October 19, 2018 recording of Donghern Kim using laboratory microscope
036	October 22, 2018 request from Respondents for an extension of time to respond to first data request
037	October 24, 2018 request from committee (Second data request) and extension of time for Respondents to respond to first data request
038	October 26, 2018 response from Respondents to committee's first data request and hard copy documents in response to part of second request for data
039	October 30, 2018 response from Respondents providing additional data in response to committee's first data request
040	November 2, 2018 response from Respondents to committee's second data request
041	November 9, 2018 request from committee for data on NIH-funded grants
042	November 16, 2018 response from Respondents to committee's request for data on NIH-funded grants
043	Respondents' list of NIH-funded grants

044	December 3, 2018 request from Lisa Cassis to HHHS ORI for a 120-day extension to complete the investigation
045	December 10, 2018 request from committee (Third request for data)
046	December 14, 2018 response from Respondents to committee's third data request
047	December 17, 2018 additional response from Respondents to committee's first data request
048	January 10, 2019 interview with Robert DiPaola
049	January 10, 2019 interview with Donghern Kim
050	Exhibits to January 10, 2019 interview with Donghern Kim
051	January 10, 2019 interview with Andrew Hitron
052	January 17, 2019 interview with Xianglin Shi
053	Exhibits to January 17, 2019 interview with Xianglin Shi and Zhou Zhang
054	January 17, 2019 interview with Zhou Zhang
055	Son, Y.O., Pratheeshkumar, P., Divya, S.P., Zhuo Zhang, Z., and Shi, X. Nuclear factor erythroid 2-related factor 2 enhances carcinogenesis by suppressing apoptosis and promoting autophagy in nickel-transformed cells. <i>Journal of Biological Chemistry</i> , 292, 8315-8330, 2017
056	Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Divya, S.P., Xu, M., Luo, J., Chen, G., Zhang, Z. and Shi, X. Antioncogenic and oncogenic properties of Nrf2 in arsenic-induced carcinogenesis. <i>Journal of Biological Chemistry</i> , 290, 27090-27100, 2015
057	Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Zhang, Z., and Shi, X. Nrf2/p62 signaling in apoptosis resistance and its role in cadmium-induced carcinogenesis. <i>Journal of Biological Chemistry</i> , 293, 15455, 2014
058	January 24, 2019 interview with Lei Wang
059	Exhibits to January 24, 2019 interview with Lei Wang
060	January 24, 2019 interview with Robert DiPaola (2)
061	Exhibit to January 24, 2019 interview with Robert DiPaola
062	January 26, 2019 additional response from Respondents to original allegation
063	January 31, 2019 interview with Donghern Kim (2)
064	Exhibits to January 31, 2019 interview with Donghern Kim
065	February 12, 2019 additional response from Respondents to original allegation
066	February 19, 2019 notification to Xianglin Shi of expansion of original allegations of research misconduct
067	February 19, 2019 notification to Zhou Zhang of expansion of original allegations of research misconduct
068	February 19, 2019 notification to Donghern Kim of expansion of original allegations of research misconduct
069	February 21, 2019 additional response from Respondents to first data request
070	February 22, 2019 request from committee that Respondents send no further data
071	March 25, 2019 request from Lisa Cassis to HHHS ORI for a 60-day extension to complete the investigation
072	April 8, 2019 notification to Xianglin Shi of further expansion of original allegations of research misconduct
073	April 8, 2019 notification to Zhou Zhang of further expansion of original allegations of research misconduct
074	April 8, 2019 notification to Donghern Kim of further expansion of original allegations of research misconduct
075	April 12, 2019 email from Xianglin Shi to Lisa Cassis in response to April 8, 2019 notification
076	April 15, 2019 email from Xianglin Shi to Lisa Cassis in further response to April 8, 2019 notification

077	April 15, 2019 email plus attachment from Bernard Pafunda (attorney for Shi and Zhang) in further response to April 8, 2019 notification
078	April 16, 2019 request from committee (Fourth data request)
079	April 18, 2019 response from Respondents requesting more time to respond to the committee's fourth data request
080	April 24, 2019 extension granted by committee giving Respondents until April 29, 2019 to respond to the committee's fourth request for data
081	April 26, 2019 request from committee for data regarding non-US funding or affiliations associated with Respondents' research or publications
082	April 29, 2019 response from Respondents to committee's fourth request for data
083	May 3, 2019 response from Respondents to committee's request for data regarding non-US funding and affiliations
084	Information from Xianglin Shi regarding non-US funding and affiliations
085	Information from Zhou Zhang regarding non-US funding and affiliations
086	Information from Donghern Kim regarding non-US funding and affiliations
087	New NIH R01 grant
088	April 2, 2019 letter from Federal Office of Research Integrity requesting that the committee investigate any non-US funding or affiliations associated with Respondents' research or publications
089	List of Respondent publications that referenced non-US funding
090	May 29, 2019 email from Kim Carter regarding disclosures of foreign funding from Respondents
091	University State Model Record Retention Schedule: Research Data
092	HHS record retention requirements
093	Section M, subsection 1.2c data
094	Section M, subsection 1.3c data
095	Section M, subsection 1.4c data
096	Section M, subsection 1.5c data
097	Section M, subsection 1.6c data
098	Section M, subsection 1.7c data
099	Section M, subsection 1.8c data
100	Section M, subsection 1.9c data
101	Section M, subsection 1.10c data
102	Section M, subsection 2.2c data
103	Section M, subsection 2.3c and 5.1c data
104	Section M, subsection 2.4c data
105	Section M, subsection 2.5c data
106	Section M, subsection 2.6c data
107	Section M, subsection 2.7c data
108	Section M, subsection 2.8c data
109	Section M, subsection 3.1c data
110	JBC guidelines: Best practices for preparing publication-quality figures
111	Section M, subsection 3.2c data
112	Section M, subsection 3.3c and 8.2c data
113	Section M, subsection 3.4c data
114	Section M, subsection 3.5c data
115	Section M, subsection 3.6c data
116	Section M, subsection 4.2c and 6.3c data
117	Section M, subsection 4.3c data

118	Section M, subsection 4.4c data
119	Section M, subsection 4.5c, 6.13c, and 7.3c data
120	Section M, subsection 5.2c data
121	Section M, subsection 5.3c data
122	Section M, subsection 6.1c data
123	Section M, subsection 6.2c data
124	Section M, subsection 6.4c data
125	Section M, subsection 6.5c data
126	Section M, subsection 6.6c data
127	Section M, subsection 6.7c data
128	Section M, subsection 6.8c data
129	Section M, subsection 6.9c data
130	Section M, subsection 6.10c data
131	Section M, subsection 6.11c data
132	Section M, subsection 6.12c and 7.2c data
133	Section M, subsection 8.1c data
134	Section M, subsection 8.3c data
135	February 21, 2019 additional data provided by Dr. Kim regarding Figure 4C in Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. <i>The Prostate</i> , 78, 390-400
136	March 1, 2019 analysis of February 21, 2019 additional data provided by Dr. Kim regarding Figure 4C in Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. <i>The Prostate</i> , 78, 390-400
137	February 12, 2019 additional material provided by Dr. Shi regarding Figure 6 in Grant Shi 3200001792
138	PubPeer comments on Respondents' retracted <i>Journal of Biological Chemistry</i> papers

L. SUMMARY STATEMENT OF FINDINGS

Executive summary of investigation into Drs. Xianglin Shi, Zhuo Zhang, and Donghern Kim

Drs. Shi, Zhang, and Kim published approximately 60 manuscripts and submitted 7 R level NIH grant proposals during the period (2012-2018) that the committee investigated. The committee examined all of these documents and identified patterns of data inconsistencies which served as the basis for the selection by the committee of 19 items (7 grants, G1-G7 (Appendices 006-012) and 12 manuscripts, M1-M12 (Appendices 013-024)) which served as a representative sample of the Respondents work. These 19 items were examined in greater detail. The investigative committee requested that the Respondents supply supporting data and documents for these 19 items. Nine distinct classes of significant departures from accepted practices of the research community were detected in the Respondents' documents. A tabulation of the committee's findings is below. A specific narrative defining these significant departures and describing each incident follows in the committee's report. The departure categories were as follows:

1. **No original data provided:** 5 figures from grants and 5 figure panels from manuscripts, each containing issues with multiple figures. *The committee defined this category as*

instances where the Respondents either did not provide any data at all or provided only PDF or PowerPoint images identical to those of the published figures.

2. **Inappropriate loading controls:** 3 figures from grants and 5 figure panels from manuscripts. *The committee defined this category as figures presented with incorrect loading controls, controls done months apart from the experiments, and/or the same controls used for multiple different experiments.*
3. **Inappropriate modification of original data:** 3 figures from grants and 3 figure panels from manuscripts. *The committee defined this category as one-dimensional stretching of original gel images, figure/lane grafting without indication, and/or cropping to remove potentially relevant data.*
4. **Inappropriate scale bars on images:** 4 figures from grants and 1 figure panel from manuscripts. *The committee defined this category as scale bars with the wrong units that could not be confirmed with metadata and/or seemed obviously wrong based on what was imaged.*
5. **Data provided did not match the published figure:** 2 figures from grants and 1 figure panel from manuscripts.
6. **Incomplete metadata provided with figure components:** 6 figures from grants and 7 figure panels from manuscripts. *The committee defined this category as images and data that lacked experimental information such as dates, exposure times, magnifications, etc.*
7. **Black images containing no pixel data:** 1 figure from grants and 2 figure panels from manuscripts. *The committee defined this category as images where, upon analysis with Photoshop, no signal was detected.*
8. **Data fabrication:** 2 figures from grants and 1 figure panel from manuscripts. *The committee defined this category as presentations of data that have been manipulated in such a way as to change the interpretation of the original data.*
9. **Falsified/fabricated data provided to the committee:** 2 instances. *The committee defined this category as presentations of data that had been falsified or altered in response to a committee query.*

After this initial investigation of grants G1-G7 (Appendices 006-012) and manuscripts M1-M12 (Appendices 013-024), and in response to the inquiry report (Appendix 002), via a letter dated April 2, 2019 from the Federal Office of Research Integrity of the Department of Health & Human Services, it was requested that the committee further examine the three retracted manuscripts from the *Journal of Biological Chemistry* below:

- R1. Son, Y.O., Pratheeshkumar, P., Divya, S.P., Zhuo Zhang, Z., and Shi, X. Nuclear factor erythroid 2-related factor 2 enhances carcinogenesis by suppressing apoptosis and promoting autophagy in nickel-transformed cells. *Journal of Biological Chemistry*, 292, 8315-8330, 2017 (Appendix 055): All figures.
- R2. Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Divya, S.P., Xu, M., Luo, J., Chen, G., Zhang, Z. and Shi, X. Antioncogenic and oncogenic properties of Nrf2 in arsenic-induced carcinogenesis. *Journal of Biological Chemistry*, 290, 27090-27100, 2015 (Appendix 056): All figures.
- R3. Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Zhang, Z., and Shi, X. Nrf2/p62 signaling in apoptosis resistance and its role in cadmium-induced carcinogenesis. *Journal of Biological Chemistry*, 293, 15455, 2014 (Appendix 057): All figures.

In these three retracted manuscripts, departures from eight (1, 2, 3, 4, 5, 6, 7, and 8) of the nine categories listed above were noted. Additionally, the committee identified that the Respondents reported incorrect percentages in flow-cytometry figures when compared to raw data (e.g., see analysis of Fig. 1F in manuscript R1). Our analysis of each of these manuscripts is found in section 10. It should be noted that during this part of the investigation, a fourth retracted paper was found in PubMed. This manuscript was not analyzed in detail but is listed in Section 10.

10. Retracted manuscripts: 3 manuscripts. *As requested by the Federal Office of Research Integrity, in this category analysis of all figures in retracted manuscripts is performed.*

CONCLUSIONS: In the course of this investigation, very little of the original data requested was provided as hard data (X-ray films, data printouts, etc.). Most data were provided in electronic formats such as PDFs of gel scans (some partial and some full gel) and PowerPoint assemblages of selected images and western blot bands. These electronic forms, in absences of hard data, are not consistent with NIH or UK regulations (Appendices 091 and 092), according to NIH policies described at https://ori.hhs.gov/sites/default/files/2018-04/5_Red_Flags_of_Research_Misconduct_scalable.pdf. Since raw data could not be produced when requested and research materials and protocols were not readily available, this constitutes a lack of transparency and departure from accepted practices. Overall, Drs. Shi and Zhang, as Principal Investigators directing the research described in the manuscripts and grants examined, have not ensured that experimental records and raw data are preserved in any systematic fashion. This lack of organization and oversight has allowed for unsupported falsified and fabricated data to be presented in grants and publications. The committee concluded that these departures from standard scientific practices were in some instances an intentional effort to deceive and in others were just careless and reckless handling of the experimental data and figure construction for grants and publications. In two instances the three Respondents (Shi, Zhang, and Kim) generated and provided falsified and fabricated documents to the committee to justify their responses to committee inquiries. It should be noted that detection of some of the significant departures from standard scientific practices described in this report required examination of the primary data. These departures would not have been identified by solely examining the published figures. Given the frequency at which issues with their data were detected in the representative sample of their work, the committee considers that this is a systemic problem that contributed to the incidents of research misconduct described in this report.

Based on review of their data and publications as well as interviews with the Respondents and observations of their laboratory protocols and procedures, the committee finds that their behavior is a significant departure from accepted practices of the research community which was committed recklessly. In specific instances (characterized in this report), the Respondents intentionally falsified and fabricated data. These allegations are supported by a preponderance of evidence, as documented in this report.

An Additional Grant Awarded During the Investigation: While completing the investigation report, the committee was made aware that Drs. Zhang and Shi had submitted and were recently awarded a new NIH R01 grant (Appendix 087). This grant was not on the list of grants provided by OSPA at the beginning of the investigation as this is a recent NIH grant application and award. During the interviews, neither Drs. Shi nor Zhang indicated that there were any more grants submitted or funded, other than the grants represented on the original list from OSPA. This new

grant has an April 1, 2019 start date and is entitled “*The role of p62 in the mechanism of Cr(VI) carcinogenesis*”. Dr. Zhang is listed as the Principal Investigator (PI) but Dr. Shi is also considered a PI, based on the budget justification.

The grant has 10 figures that include micrographs, western blots, and images of animals. The load controls for Figs. 1, 5, 7A, 8 and 9 are suspect and would require original data to verify. It should be noted that Fig. 7A presents the same data as is discussed in Section M subsection 2.3 of this report. Fig. 8 presents the same data as is discussed in the committee’s report in Section M subsection 2.2. This type of departure from professional standards has been noted in other work from Drs. Zhang and Shi; see further examples in Section M subsection 2. Upon inspection, the scale bars in Figs. 2, 3C, and 10 are not correct. In particular, the units on the scale bars in Fig. 10A are incorrect. It should be noted that Fig. 10B presents the same data as is discussed in the committee’s report in Section M subsection 4.1 below. This type of departure from professional standards has been noted in other work from Drs. Zhang and Shi; see further examples in Section M subsection 4.

The committee concluded that the noted figures presented in this grant are representative of the types of departures from professional standards detected in other work from the Respondents. The conclusions made in Sections M subsection 2 and M subsection 4 are equally applicable to this grant.

Funding From Foreign Sources: The Federal Office of Research Integrity requested that the committee investigate the possibility of non-US affiliation or funding associated with the Respondents’ research and/or publications (Appendix 088). The committee generated a list of publications and grants where foreign funding had been indicated (Appendix 089). This list was supplied to the Respondents on April 26, 2019 and a request made that the Respondents explain any foreign funding or affiliations (Appendix 081), and the Respondents replied on May 3, 2019, indicating that in each case, to the extent there was a foreign funding or affiliation, it was that of a co-author and not one of the Respondents ((Appendices 083 (email), 084 (for Shi), 085 (for Zhang), and 086 (for Kim)). Further, the list was provided to the Office of Sponsored Research Administration (OSPA) for further analysis. The OSPA director, Kim Carter, stated that “A review of the NIH awards included in the publication list provided below did not reveal any relevant disclosures of foreign funding or affiliations for Drs. Shi and Zhang. Please note that many of the award files were not available in OSPA e-files because they had been destroyed in accordance with the University’s records retention policy. Information available in NIH Commons was used to review some of the awards.” (Appendix 090). Non-US funding disclosure is not a research misconduct issue as it does not qualify as fabrication, falsification or plagiarism; however, failure to disclose foreign ties on an NIH grant application violates long-standing departmental (Health and Human Services) rules since such disclosures are part of a broader NIH requirement that scientists must declare “all financial resources”.

M. DETAILED ANALYSIS OF FINDINGS IN THE 10 SIGNIFICANT DEPARTURE CATEGORIES IDENTIFIED

1. No original data provided:

Description: The committee defined this category as instances where the Respondents either did not provide any data at all or provided only PDF or PowerPoint images identical to those of the published figures.

Overall summary: 5 figures from grants and 5 manuscripts with multiple figure had no original data provided (sections 1.1-1.10). Only digital data with no or minimal labeling was provided.

Overall summary of interviews and relevant comments: During the interviews, the committee asked each Respondent, as well as two other laboratory personnel (Drs. Hitron and Wang), a series of questions about how original data were retained and reviewed. Original western blot films were generally converted into digital images and used to prepare PowerPoint presentations. Both Dr. Shi and Dr. Zhang stated that they considered these PowerPoints to be original data (Appendices 052 (Shi interview pages 16, 17, 99, 101) and 054 (Zhang interview, page 86)). Dr. Zhang stated that she was unclear about what “original” films meant (Appendix 054 (Zhang interview, page 85)). She stated that she believed that scanned versions of data are the same as the original film (Appendix 054 (Zhang interview, page 87)). Shi further claimed that he was told by other people at UK that they kept their records in PowerPoint, and that keeping records in PowerPoint was sufficient for NIH record retention purposes (Appendix 052 (Shi interview, pages 16-17, 101)).

Overall conclusions: Based on the interviews, the committee’s observations of the state of the laboratory records and the Respondents’ inability to provide original data in only one instance, there appeared to be no overall policy for storage, annotation, or organization of original experimental data. In addition, there was no standardization of notebook keeping or relating laboratory notes to original data. These deficiencies resulted in an inability to find, produce, and confirm the original data as they relate to a specific experiment in a grant or publication. In most cases, the electronic data provided was connected to a particular experiment by simply matching it with the published figure rather than using experimental notes. Thus, it is impossible to verify the experimental details associated with the data that went into grants and publications. Despite explicit instructions from the committee for the Respondents to produce raw images/data in response to the committee’s requests (Appendices 034 and 037), the Respondents produced only electronic versions of that data. When the committee further specified, in its third request for data, that the Respondents must produce original, hard copy data (“This documentation must include hard copies and computer file paths for all protocols, raw data, final figures, and any intermediates leading to the final submitted figures. Scans of data are not sufficient. All raw data and protocols should be dated, and any electronic data or files should contain the original metadata that documents when the data were recorded and how.” (Appendix 045)), the Respondents continued to provide only electronic versions. The Respondents only produced some original data in response to the committee’s fourth request for data (Appendix 078).

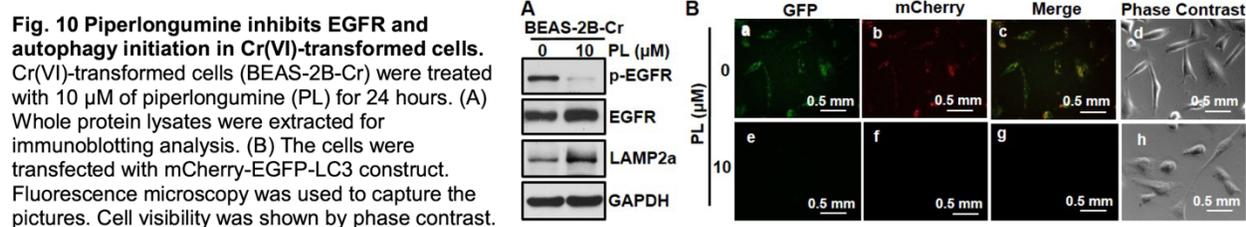
In some cases, Drs. Shi and Zhang stated that the original data were unavailable because they were out of the country with a former lab member (Appendix 093). These deficiencies are clear deviations from the data storage guidelines of UK (Appendix 091) and federal agencies (Appendix 092).

1. Specific examples for “no original data provided”:

1.1. G1: Grant Shi_3210000529 = 3R01ES025515-03S1 (02/01/2017-01/30/2020) (Appendix 006, pages 11-133): Fig. 10Bg

1.1a. *This figure was originally provided to the committee at the start of the investigation.*

1.1b. *Original figure from Appendix 006 page 63 of the research strategy section of the grant.*



1.1c. *Data provided.* See inquiry report (Appendix 006, pages 11-133). In addition, Dr. Shi, after the interviews, provided some file paths for Fig. 10 (Appendix 062).

1.1d. *Our analysis.* The committee concurs with the inquiry report conclusions (page 139 in Appendix 006). The committee was able to find the Differential Interference Contrast (DIC) image for Fig. 10H on the sequestered hard drives. However, the committee found no other fluorescence images that were taken at the same time, as would be expected if panels E-H represented the same experiment. Images for Figs. 10A-D were found and had consistent time stamps.

1.1e. *Relevant interview questions/comments.* See inquiry report (Appendix 006, page 52 of Exhibit C in Appendix K). Shi is not able to explain why there is no data in 10B (G) (Appendix 052 (Shi interview, page 41)). Zhang does not know why there is no data in Figure 10B (G). She says Kim took the pictures (Appendix 054 (Zhang interview, pages 44-45)). Kim did the experiment and Zhang and another lab member asked him if the cells were still alive. Zhang asked Kim to repeat the experiment. He took a picture of six fields (Appendix 063 (Kim interview II, page 14)). Kim used Fig. 6 as a template to make Fig. 10. He overlaid Fig. 10 on Fig. 6 and the bottom slide on both figures is blank (Appendix 063 (Kim interview II, page 15)). All six fields were blank. They all had the same exposure and contrast. The exposure time was the same over panels A and E, G and F (Appendix 063 (Kim interview II, page 15)).

1.1f. *Conclusion for this specific example.* The Respondents could not verify the validity of Fig. 10E-G. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

1.2. G2: Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Fig. 9

1.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

1.2b. *Original figure from Appendix 007 page 4 of the research strategy section of the grant.*



Fig. 9 Luteolin increases Nrf2 expression in normal cells but decreases constitutive Nrf2 expression in Cr(VI)-transformed cells. Normal parent BEAS-2B cells and Cr(VI)-transformed BEAS-2B cells (B2B-Cr) were treated with luteolin (10 μ M) for 24 hours. Cells were harvested and whole protein lysates were extracted. Nrf2 expression was examined by immunoblotting.

- 1.2c. *Data provided.* From the Respondents (Appendix 093): “This work (Fig. 9) was done by Dr. Poyil Pratheeshkumar. He was a postdoc from 2011-2016. He is now at Human Cancer Genomic Research, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. Dr. Pratheeshkumar provided the PPT data above. Dr. Shi asked him for original data. His answer is that he stored the data in his home country, India and cannot access it anytime soon.”
- 1.2d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).
- 1.2e. *Relevant interview questions/comments.* Dr. Shi stated that he was not able to find the original data for Figures 9 and 10 in the lab. He had to ask Dr. Pratheeshkumar for it. Electronic versions of the data were sent to Dr. Shi by Dr. Pratheeshkumar (Appendix 052 (Shi interview, pages 64-66)).
- 1.2f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 1.3. G2: Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Fig. 10
- 1.3a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 1.3b. *Original figure from Appendix 007 page 4 of the research strategy section of the grant.*

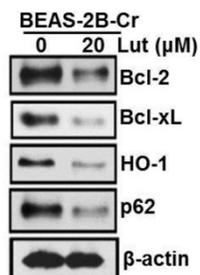


Fig. 10 Luteolin inhibits constitutive expressions of Bcl-2, Bcl-xL, HO-1, and p62 in Cr(VI)-transformed cells. Cr(VI)-transformed BEAS-2B cells (BEAS-2B-Cr) cells were treated with luteolin (20 μ M) for 24 hours. Cells were harvested and whole protein lysates were extracted. Expressions of Bcl-2, Bcl-xL, HO-1, and p62 were examined by immunoblotting.

- 1.3c. *Data provided.* From the Respondents (Appendix 094): “Dr. Pratheeshkumar provided the results in PPT. He was a postdoc from 2011-2016. He is now at Human Cancer Genomic

Research, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. Dr. Shi asked him for original data. His answer is that he stored the data in his home country, India and cannot access it anytime soon.”

- 1.3d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).
- 1.3e. *Relevant interview questions/comments.* Dr. Shi stated that he was not able to find the original data for Figure 10 in the lab. He had to ask Dr. Pratheeskumar for it. Electronic versions of the data were sent to Dr. Shi by Dr. Pratheeskumar (Appendix 052 (Shi interview, pages 64-66)).
- 1.3f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 1.4. G4: Grant Zhang_3048111797 = 1R01ES021771-01A1 (08/01/2014-04/30/2019) (Appendix 009): Fig. 3
- 1.4a. *Date requested:* October 18, 2018 by the committee via email.
Date received: October 26, 2018 by Respondents via flash drive.
- 1.4b. *Original figure from Appendix 009 page 3 of the research strategy section of the grant.*

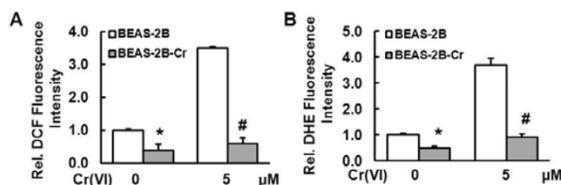
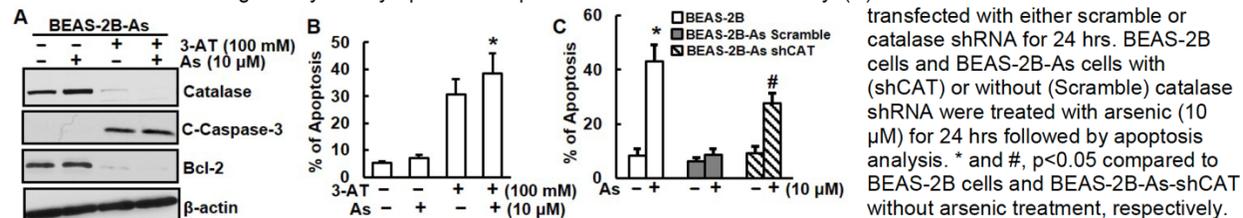


Fig. 3 Reduced ROS generation in Cr(VI)-transformed cells. Both non-transformed (BEAS-2B) and Cr(VI)-transformed BEAS-2B (BEAS-2B-Cr) cells were treated with 5 μM Cr(VI) for 6 hr and analyzed for DHE and DCF fluorescence intensities by flow cytometry. Data are expressed as mean±SD (n=6). * and #, p < 0.05 compared to passage-matched BEAS-2B cells without treatment and with 5 μM Cr(VI) treatment, respectively.

- 1.4c. *Data provided.* From the Respondents (Appendix 095): “Dr. Lijuan Sun, a postdoc fellow from 2010-2012, performed this experiment. We are unable to locate the raw data at this time. We are still trying to find them.”
- 1.4d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).
- 1.4e. *Relevant interview questions/comments.* Not applicable.
- 1.4f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 1.5. G5: Grant Zhang_3200001472 = 5R01ES028321-02 (09/01/2017-07/31/2022) (Appendix 010): Fig. 10
- 1.5a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 1.5b. *Original figure from Appendix 010 page 32 of the grant.*

Fig. 10 Inhibition of catalase activity restores apoptosis in arsenic-transformed cells. (A) and (B) Arsenic-transformed cells (BEAS-2B-As) were treated with different conditions as indicated for 24 hrs. A, Expressions of catalase, C-caspase-3, and Bcl-2 were examined using immunoblotting. The results are representative of three independent experiments. B, The percentage of apoptotic cells was measured using flow cytometry. *p<0.05 compared to arsenic-treated cells only. (C) Arsenic-transformed cells were



- 1.5c. *Data provided.* From the Respondents (Appendix 096): “Dr. Amit Budhaja, a graduate student (2007-2012) and a postdoc (2012-2013). He did not save the whole-film original images.”
- 1.5d. *Our analysis.* The Respondents were unable to provide the original data but did provide data from a repeated experiment that confirmed the results for Fig. 10A. They also provided screenshots of the spreadsheets for the data in Figs. 10B and 10C, whose veracity the committee could not confirm. (See Appendix 096)
- 1.5e. *Relevant interview questions/comments.* Not applicable.
- 1.5f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 1.6. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 3G
- 1.6a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 1.6b. *Original figure from Appendix 013 page 6.*

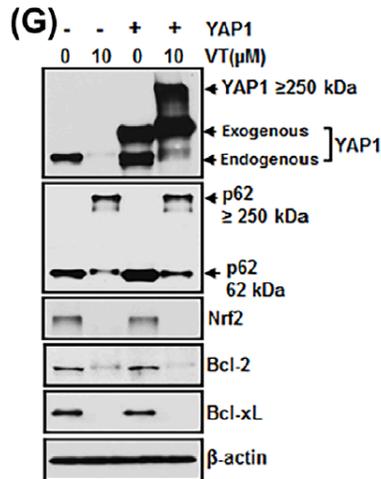


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μM hydroxychloroquine (HCQ), 10 μM Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μM Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μM and 10 μM Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μM 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean ± SE ($n = 6$). * and #, $P < 0.05$ compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μM of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

- 1.6c. *Data provided.* From the Respondents (Appendix 097): “I (Lei Wang) have no ideal for why I cannot find the original data in Fig. 3G (Nrf2 and Bcl-2). There are kind of reasons to make it lost: wrongly kept files, wrongly delete, computer update lost, and copy lost et al. I can just find the very original one is the PPT file which was attached here.”
- 1.6d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).
- 1.6e. *Relevant interview questions/comments.* Not applicable.
- 1.6f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 1.7. M2: Son YO, Pratheeshkumar P, Wang Y, Kim D, Zhang Z, and Shi X. (2017). Protection from Cr(VI)-induced malignant cell transformation and tumorigenesis of Cr(VI)-

transformed cells by luteolin through Nrf2 signaling. *Toxicology and Applied Pharmacology*, 331, 24-32 (Appendix 014): All figures

1.7a. *Date requested*: October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

1.7b. *Original figure from Appendix 014*.

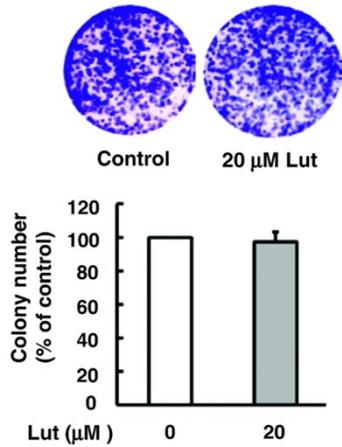


Fig. 1. Luteolin does not exhibit any cytotoxic property. Cell viability was measured by clonogenic assay. BEAS-2B cells were treated without or with luteolin (20 μM) for 48 h, reseeded and cultured for an additional 2 weeks and stained with crystal violet. The data are expressed as the mean ± SE (n = 6). Lut; luteolin.

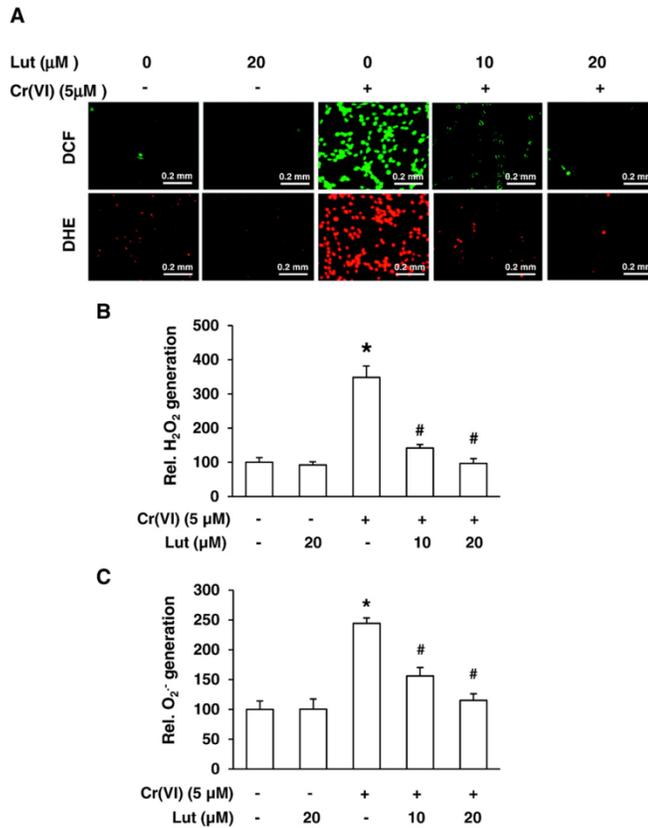


Fig. 2. Luteolin decreases the levels of H₂O₂ and O₂⁻ generated by Cr(VI). BEAS-2B cells were exposed to Cr(VI) (0 or 5 μM) with or without luteolin (0, 10, 20 μM) for 12 h and then were labeled with DHE (5 μM) or DCFDA (10 μM). (A) Images were taken with fluorescence microscopy and fluorescent intensity determined by flow cytometry. Micrographs represent one of six replicates from each treatment group. Graphic representation of fluorescence intensity for (B) H₂O₂ and (C) O₂⁻ with *p ≤ 0.05 compared to control, #p ≤ 0.05 compared to Cr(VI) alone. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells; Lut; luteolin.

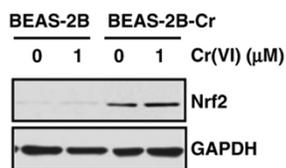


Fig. 3. Increased levels of Nrf2 in Cr(VI)-transformed cells. Both normal parental and Cr(VI)-transformed BEAS-2B cells were treated with and without 1 μM of Cr(VI) for 24 h. The cells were harvested and whole protein lysates was obtained. The expression of Nrf2 was analyzed by immunoblot. Presented results are representative of three independent experiments. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells.

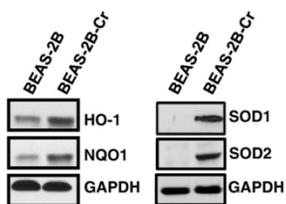


Fig. 4. Increased levels of antioxidant proteins in Cr(VI)-transformed cells. Whole cell lysates were extracted from Cr(VI)-transformed cells and passage-matched normal cells (BEAS-2B) for immunoblot analysis. Presented results are representative of three independent experiments. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells.

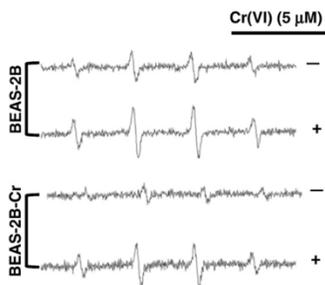


Fig. 5. Reduced ROS generation in Cr(VI)-transformed cells as measured by electron spin resonance spin trapping. Both normal parental BEAS-2B cells and Cr(VI)-transformed BEAS-2B cells were treated with and without 5 μM Cr(VI) for 12 h. The oxygen radical generations were measured by ESR spin trapping using 5,5-Dimethyl-1-pyrroline *N*-oxide DMPO. The results are representative of three independent experiments. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells.

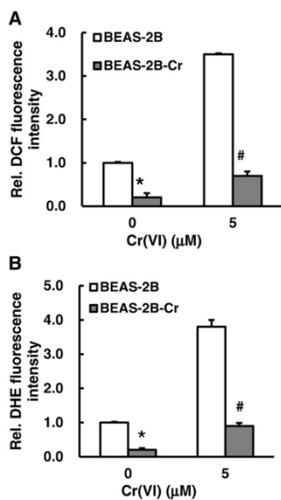


Fig. 6. Reduced ROS generation in Cr(VI)-transformed cells as measured by fluorescence. Both normal parental BEAS-2B cells and Cr(VI)-transformed BEAS-2B cells were treated with and without 5 μM Cr(VI) for 12 h and analyzed for DCF (A) and DHE (B) fluorescence intensities by flow cytometry as described in previously. Data are mean \pm SE (n = 6). * and # indicates $p < 0.05$ compared to normal parent BEAS-2B cells without and with 5 μM Cr(VI) treatment, respectively. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells.

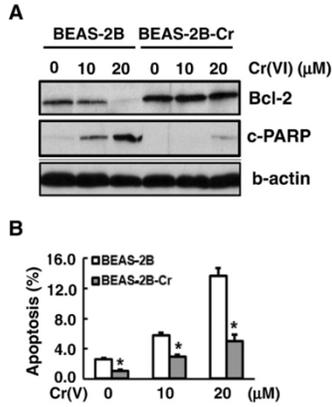


Fig. 7. Acquired apoptosis resistance in Cr(VI)-transformed BEAS-2B cells. (A) Both BEAS-2B and Cr(VI)-transformed cells were treated with different concentrations of Cr(VI) for 24 h. The whole cell lysates were collected to examine expressions of Bcl-2 and cleaved PARP (c-PARP) by immunoblot analysis. (B) Apoptosis was measured using FITC-conjugated Annexin V/PI staining by flow cytometry. Both BEAS-2B and Cr(VI)-transformed cells were seeded into 6-well culture plates (1×10^5 cells/well). The cells were treated with different concentrations of Cr(VI). After 24 h, 5 μl of fluorescein isothiocyanate (FITC)-labeled PI solution (10 μg/ml) was added into the cells. The percentage of apoptotic cells was measured using flow cytometry. Data are mean \pm SE (n = 6). * indicates a $p < 0.05$, compared to normal BEAS-2B cells. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells.

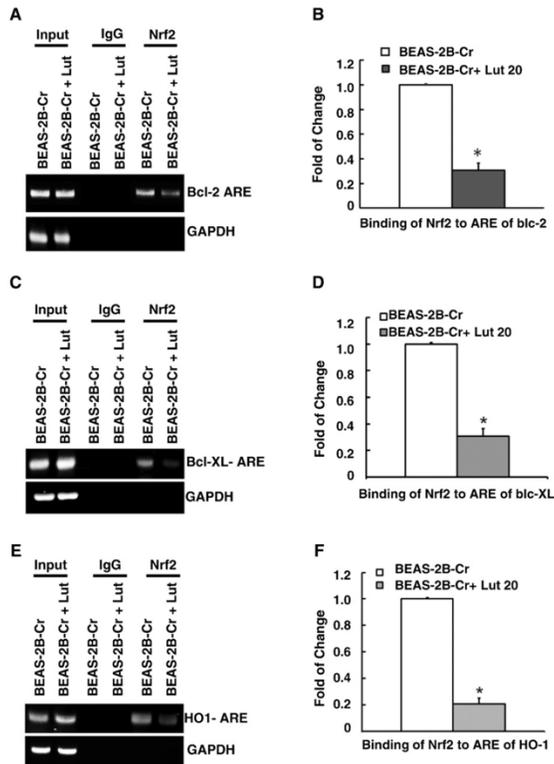


Fig. 8. Luteolin inhibits binding of Nrf2 to ARE sites on the *Bcl-2*, *Bcl-XL*, and *HO-1* gene promoters in Cr(VI)-transformed cells using ChIP assay. Cr(VI)-transformed cells in the presence or absence of luteolin (20 μM) were fixed with formaldehyde and cross-linked. The chromatin was sheared and immunoprecipitated with anti-Nrf2 antibody or control IgG. Binding of Nrf2 to Bcl-2, Bcl-XL, and HO-1 promoters was analyzed by PCR using specific primers for ARE regions of their promoters (A, C, and E). ChIP and quantitative RT-PCR analysis was performed. The immunoprecipitated DNA was normalized to the input levels and plotted (B, D, and F). Data are mean \pm SE (n = 6). * indicates a $p < 0.05$, compared to levels obtained without luteolin. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells; Lut; luteolin.

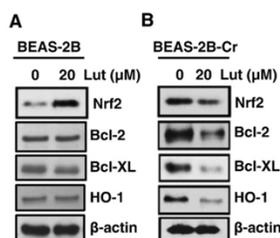


Fig. 9. Luteolin increases inducible Nrf2 expression in normal cells and decreases constitutive Nrf2 expression in Cr(VI)-transformed cells. Normal parental BEAS-2B cells (A) and Cr(VI)-transformed BEAS-2B cells (B) were treated with luteolin (10 μM) for 24 h. Cells were harvested and whole protein lysates were extracted. The expression levels of Nrf2 and its target proteins were examined by immunoblot analysis. Presented results are representative of three independent experiments. BEAS-2B-Cr; Cr(VI)-transformed BEAS-2B cells; Lut; luteolin.

1.7c. *Data provided.* From the Respondents (for each figure):

For Figs. 1, 2, and 9 (Appendix 098): “Dr. Poyil Pratheeshkumar provided the results in PPT. He was a postdoc from 2011-2016. He is now at Human Cancer Genomic Research, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. Dr. Shi asked him for original data. His answer is that he stored the data in his home country, India and cannot access it any time soon.”

For Fig. 3 (Appendix 098): “This work (3) was done by Dr. Young-OK Son. He was a postdoc from 2008-2014. He is now at National Creative Research Initiatives Center for Osteoarthritis Pathogenesis and School of Life Sciences, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea.”

For Fig. 4 (Appendix 098): “Dr. Donghern Kim, a postdoc, generated the data in the left and provided the results in PPT. Dr. Xin Wang, a postdoc (2008-2013), generated the data and provided the results in PPT. The original data has not been located. We are still trying.”

Screenshot provided for Fig. 5 (Appendix 098).

For Figs. 6 and 7 (Appendix 098): “The work was done by Dr. Lijuan Sun, a postdoc from 2010 to 2012. We are still trying to locate the raw data.”

Data were provided for Fig. 8 (Appendix 095). Full agarose gels for panels A, C and E, and Excel spreadsheet data for panels B, D and F were provided (Appendix 098). The data provided were in an electronic format only.

1.7d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).

1.7e. *Relevant interview questions/comments.* Not applicable.

1.7f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 1.8. M4: Gao, N., Cheng, S., Budhraja, A., Liu, E.H., Chen, J., Chen, D., Yang, Z., Luo, J., Shi, X., and Zhang, Z. 3,3'-Diindolylmethane exhibits antileukemic activity in vitro and in vivo through a Akt-dependent process. *PLoS One*. 7, e31783, 2012 (Appendix 016): All figures
- 1.8a. Date requested: October 24, 2018 by the committee via email.
Date received: November 2, 2018 by Respondents via flash drive.
- 1.8b. Original figure from Appendix 016 pages 2-7.

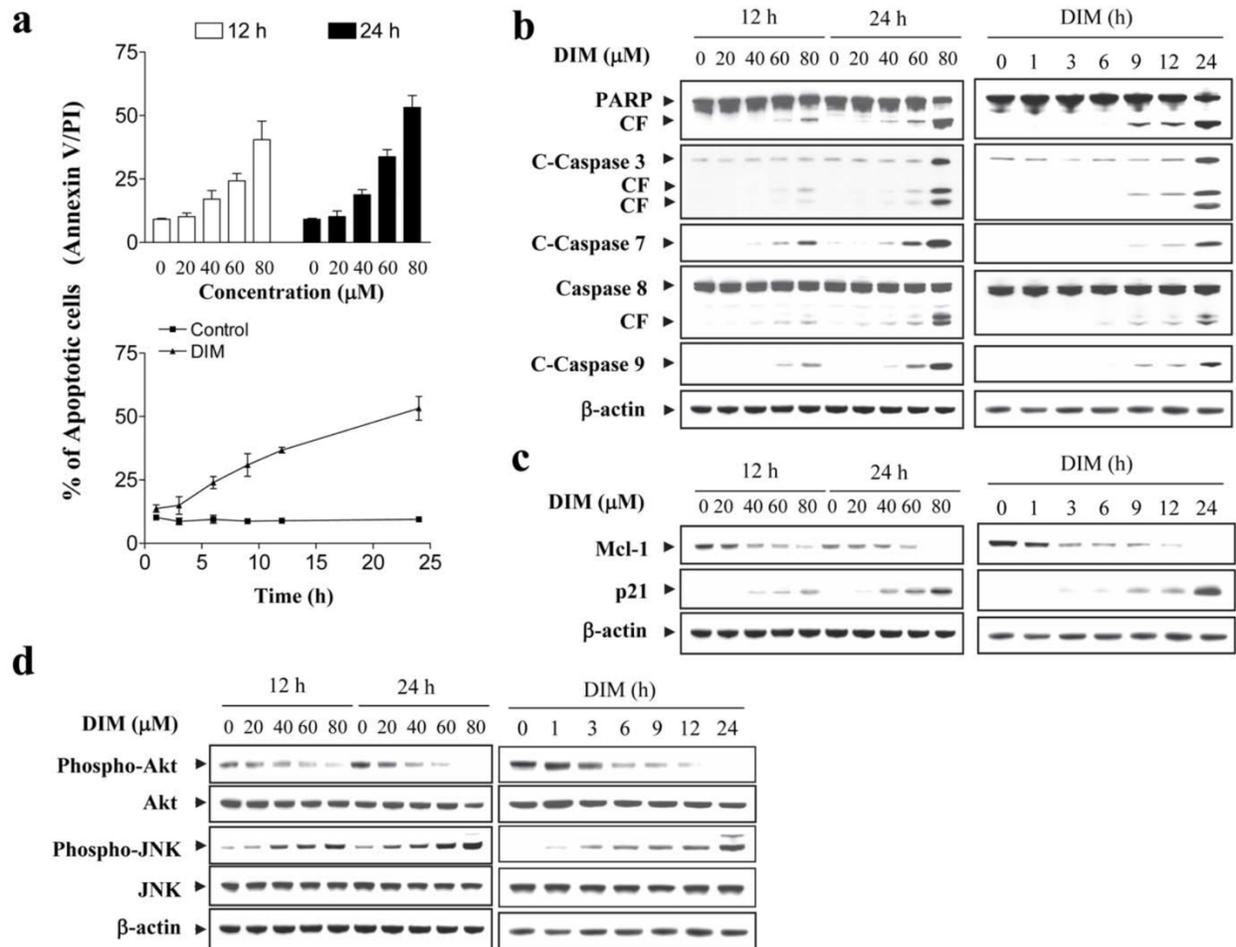


Figure 1. DIM induces apoptosis, caspase activation, downregulation of Mcl-1, upregulation of p21, inactivation of Akt, and activation of JNK in U937 human leukemia cells in dose- and time-dependent manners. U937 cells were treated with various concentrations of DIM as indicated for 12 h and 24 h or treated with 80 μM DIM for 1, 3, 6, 9, 12, and 24 h. (a) Cells were washed twice with PBS and stained with Annexin V/propidium iodide (PI), and apoptosis was determined using flow cytometry. Both early apoptotic (Annexin V-positive, PI-negative) and late apoptotic (Annexin V-positive and PI-positive) cells were included in cell death determinations. The values obtained from annexin V/PI assays represent the mean \pm SD for three separate experiments. (b–d) Total cellular extracts were prepared and subjected to Western blot assay using antibodies as indicated.

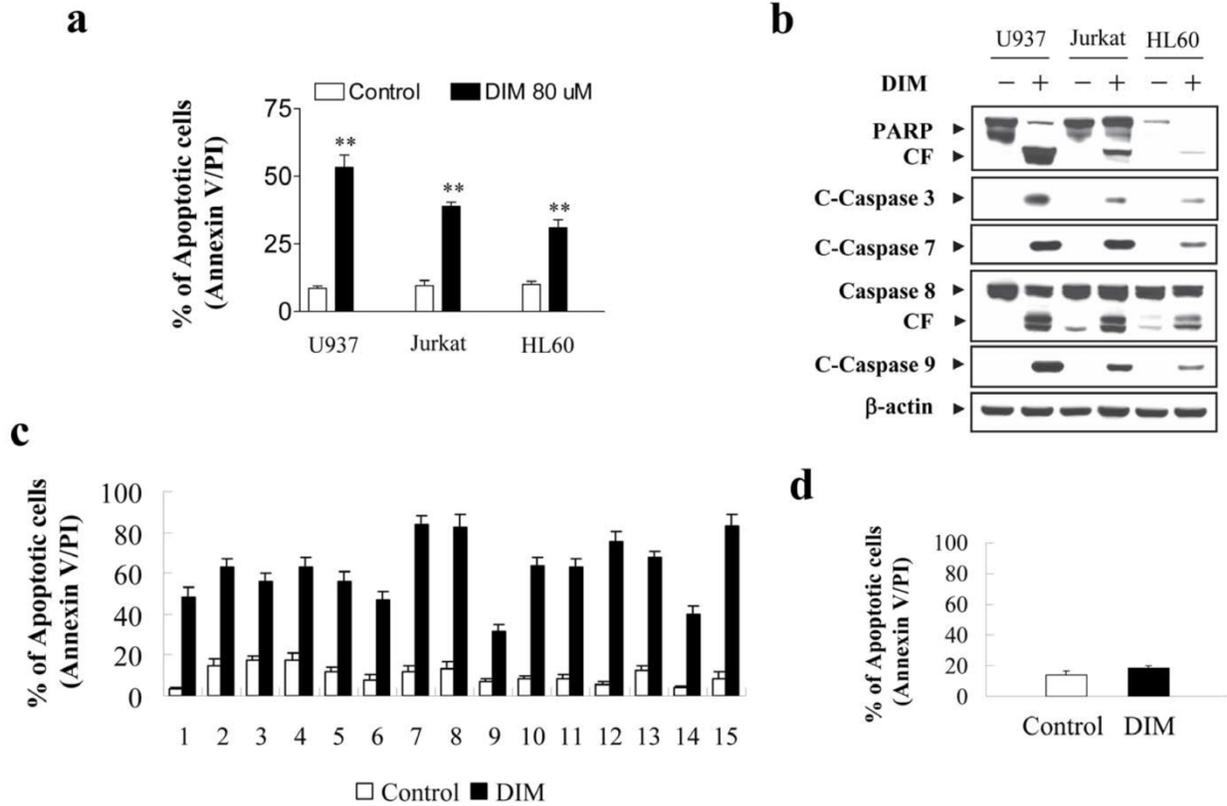


Figure 2. DIM induces apoptosis in U937, Jurkat, and HL-60 cells, and in AML blast samples, but not in normal bone marrow mononuclear cells. (a) U937, Jurkat, and HL-60 cells were treated with 80 μ M DIM for 24 h, after which apoptosis was determined by annexin V/PI staining and flow cytometry. ** Values for cells treated with DIM were significantly increased compared to values in control cells by Student's t-test, $p < 0.01$. (b) Total cellular extracts were prepared and subjected to Western blot analysis using antibodies as indicated. (c-d) Blasts from 15 patients with AML and normal bone marrow mononuclear cells were treated with 80 μ M DIM for 24 hours, apoptosis was determined by annexin V/PI staining and flow cytometry.

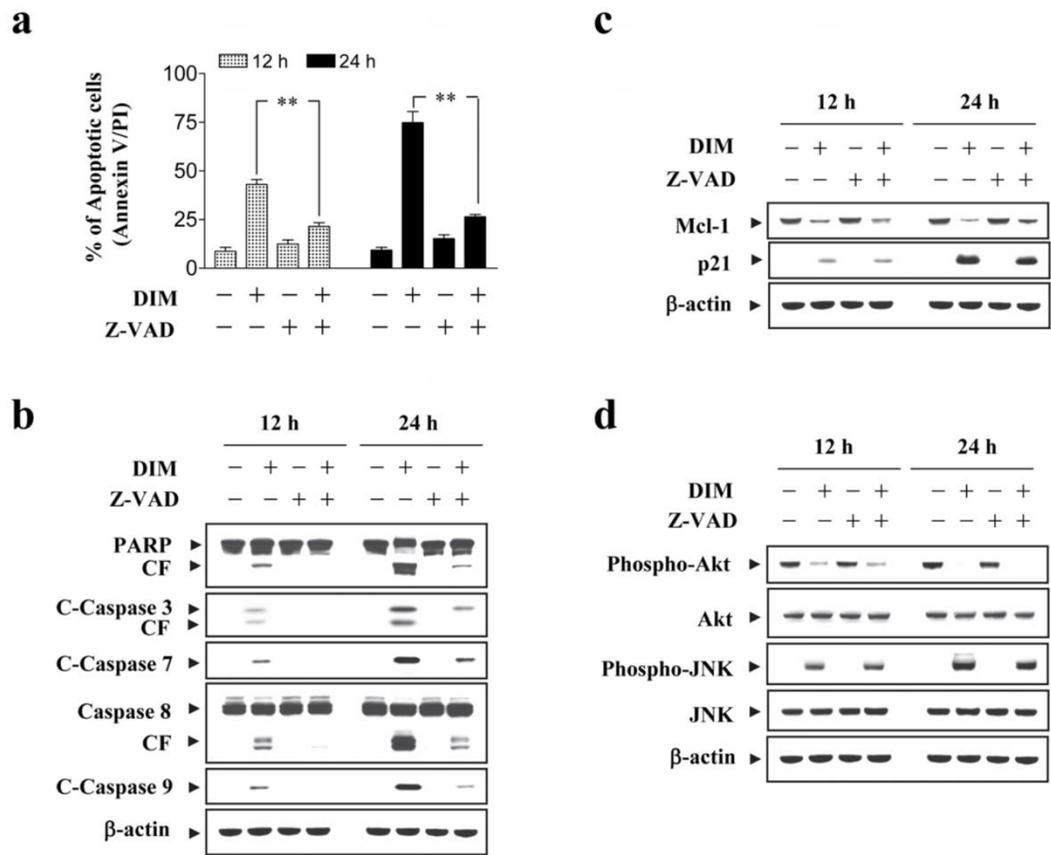


Figure 3. Effects of inhibition of caspases by Z-VAD-FMK on apoptosis, expression of Mcl-1 and p21, and phosphorylation of Akt and JNK. U937 cells were pretreated with the caspase inhibitor Z-VAD-FMK (20 μ M) for 1 h, followed by treatment with 80 μ M DIM for 12 h and 24 h. (a) Cells were stained with Annexin V/PI, and apoptosis was determined using flow cytometry. **Values for cells treated with DIM and Z-VAD-FMK were significantly reduced compared to values obtained for DIM alone by Student's t-test, $p < 0.01$. (b-d) Total protein extracts were prepared and subjected to Western blot assay using antibodies as indicated.

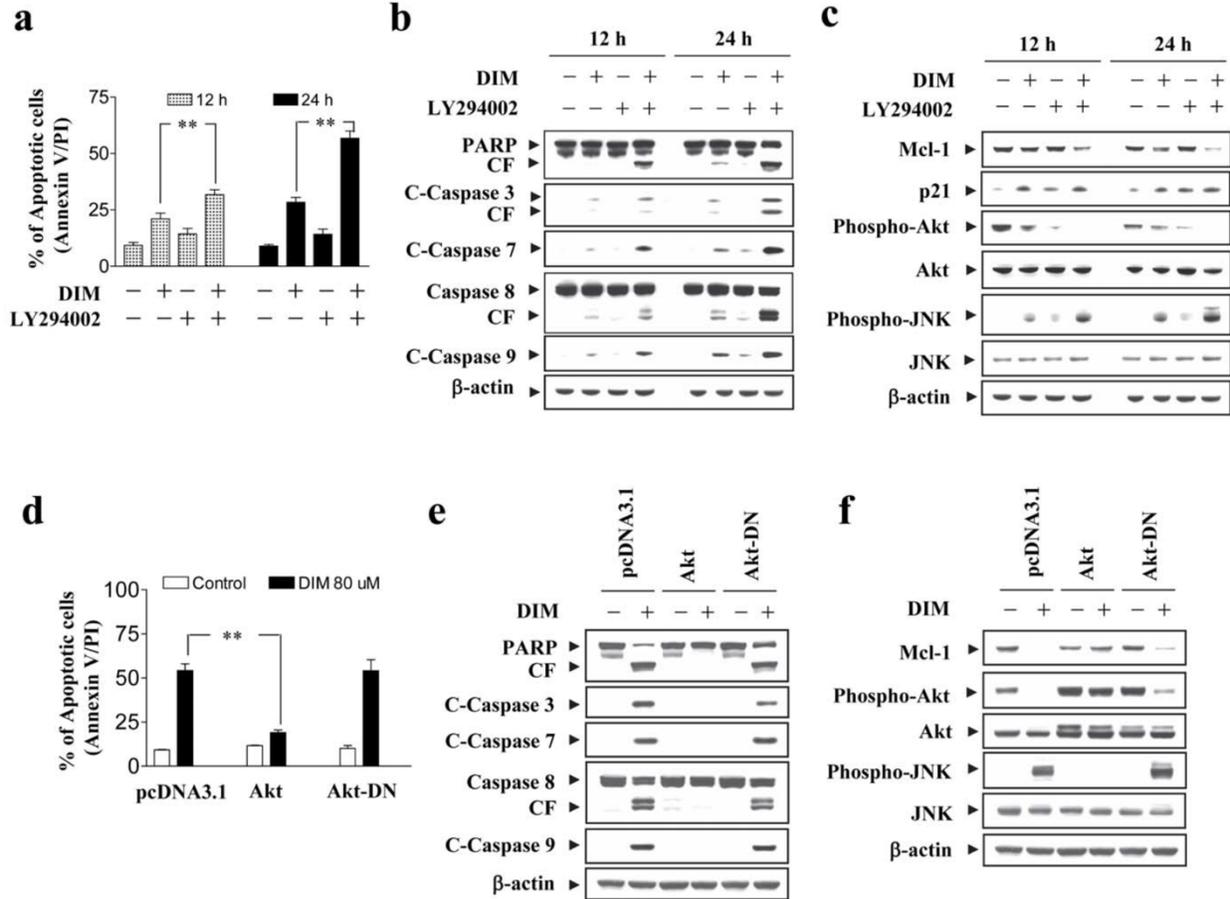


Figure 4. Effects of PI3K inhibitor, LY294002 (LY) and genetic activation of Akt on apoptosis induced by DIM. U937 cells were pretreated with 20 μ M of LY for 1 h, followed by the addition of 40 μ M of DIM for 24 h. (a) Cells were stained with Annexin V/PI, and apoptosis was determined using flow cytometry. **Values for cells treated with DIM and LY were significantly greater than those for cells treated with DIM alone by Student's *t*-test; $p < 0.01$. (b–c) Total cellular extracts were prepared and subjected to Western blot analysis using antibodies as indicated. (d) U937 cells were stably transfected with an empty vector (pcDNA3.1), Akt-CA, and Akt-DN. Cells were treated with 80 μ M of DIM for 24 h, after which apoptosis was analysed using Annexin V/PI assay. **Values for Akt-CA cells treated with DIM were significantly decreased compared to those for pcDNA3.1 cells by Student's *t*-test, $p < 0.01$. (e–f) Total cellular extracts were prepared and subjected to Western blot analysis using antibodies as indicated.

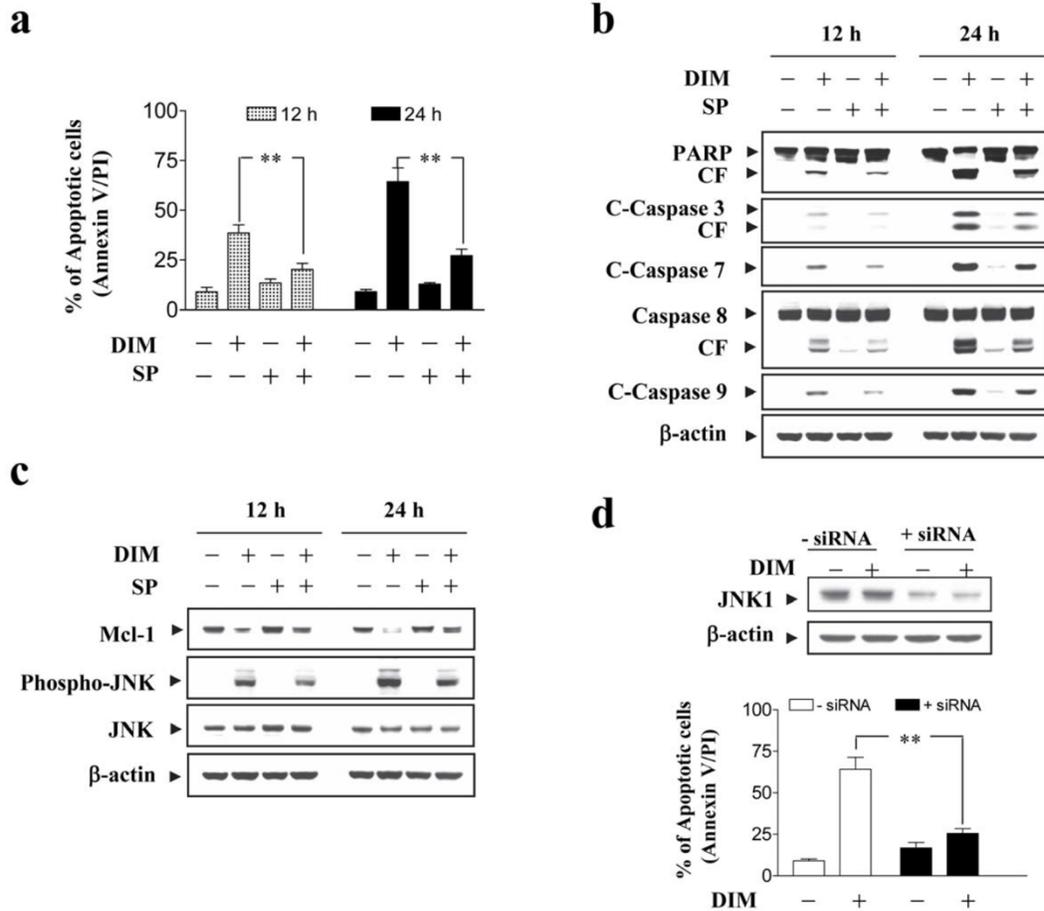


Figure 5. Inhibition of JNK significantly protect cells from DIM-induced apoptosis. U937 cells were pretreated with 10 μ M of JNK inhibitor, SP600125 (SP), for 1 h, followed by the addition of 80 μ M of DIM for 24 h. (a) Cells were stained with Annexin V/PI, and apoptosis was determined using flow cytometry. **Values for cells treated with DIM and SP were significantly less than those obtained for cells treated with DIM alone by Student's t-test, $p < 0.01$. (b–c) Total cellular extracts were prepared and subjected to Western blot assay using antibodies as indicated. (d) U937 cells were transfected with JNK1 siRNA oligonucleotides or controls and incubated for 24 h at 37°C, after which cells were treated with 80 μ M of DIM for 24 h. Apoptosis was determined using the Annexin V/PI assay. **Values for cells treated with DIM after transfection with JNK1 siRNA were significantly decreased compared to those for control cells treated with DIM by Student's t-test; $p < 0.01$. Total cellular extracts were prepared and subjected to Western blot analysis using antibodies against JNK1.

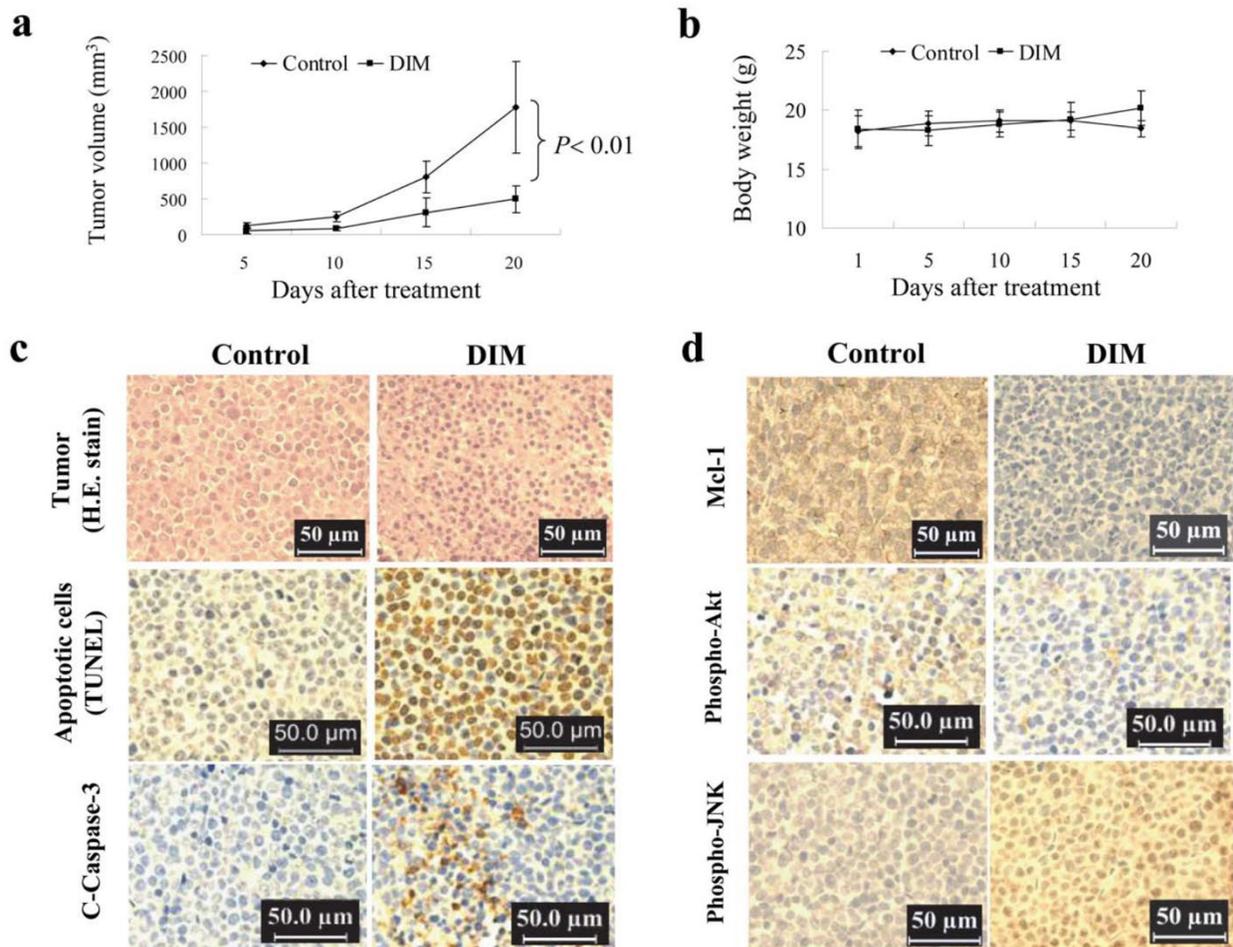


Figure 6. *In vivo* antileukemic activity of DIM in U937 xenografts. 20 NOD/SCID mice were inoculated with U937 cells (2×10^6 cells/mouse, i.p.) and randomly divided into two groups (10/group) for treatment with DIM (50 mg/kg, i.p., daily, five times per week) or with vehicle control solvent. (a) Average tumor volume in vehicle control mice and mice treated with 50 mg/kg DIM. $P < 0.01$, significantly different compared with vehicle control by Student's *t*-test. (b) body weight changes of mice during the 20 days of study. (c) At the 20 days after DIM treatment, the tumors were excised and subjected to H&E staining for determination of pathological evaluation, TUNEL assay for determination of apoptosis, and immunohistochemical staining to determine Cleavage-caspase-3 immunoreactivity. Original magnification $\times 400$. (d) After treatment with DIM, tumor tissues were sectioned and subjected to immunohistochemistry using antibodies as indicated.

1.8c. *Data provided.* From the Respondents (Appendix 099): “Dr. Ning Gao is the first author and corresponding author in this publication. Dr. Gao was joint faculty in the Graduate Center for Toxicology at that time. The data were generated in Dr. Gao’s laboratory in China. See the email attached [in Appendix 099] from journal editorial office indicated that Drs. Shi and Zhang were co-authored. We are unable to provide raw data.”

1.8d. *Our analysis.* Since the funding listed on this manuscript included the NIH (grants ES015375 to Dr. Shi, and ES019249 to Dr. Zhang), the original data should have been retained based on UK (Appendix 091) and federal guidelines (Appendix 092). The other funding listed on this manuscript was from National Natural Science Foundation of China (grant 30971288) without specification of who the PI was; however, Dr. Shi and Zhang both reported that this funding was supposedly to their co-author Dr. Ning Gao (Appendix 083).

1.8e. *Relevant interview questions/comments.* Not applicable.

- 1.8f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 1.9. M3: Pratheeshkumar P, Son YO, Divya SP, Wang L, Turcios L, Roy RV, Hitron JA, Kim D, Dai J, Asha P, Zhang Z, and Shi X. (2017). Quercetin inhibits Cr(VI)-induced malignant cell transformation by targeting miR-21-PDCD4 signaling pathway. *Oncotarget*, 8, 52118-52131 (Appendix 015): All figures requested
- 1.9a. *Date requested:* December 10, 2018 by the committee *via* email.
Date received: December 14, 2018 by Respondents *via* flash drive.
- 1.9b. *Original figures from Appendix 015 pages 52119-52125.*

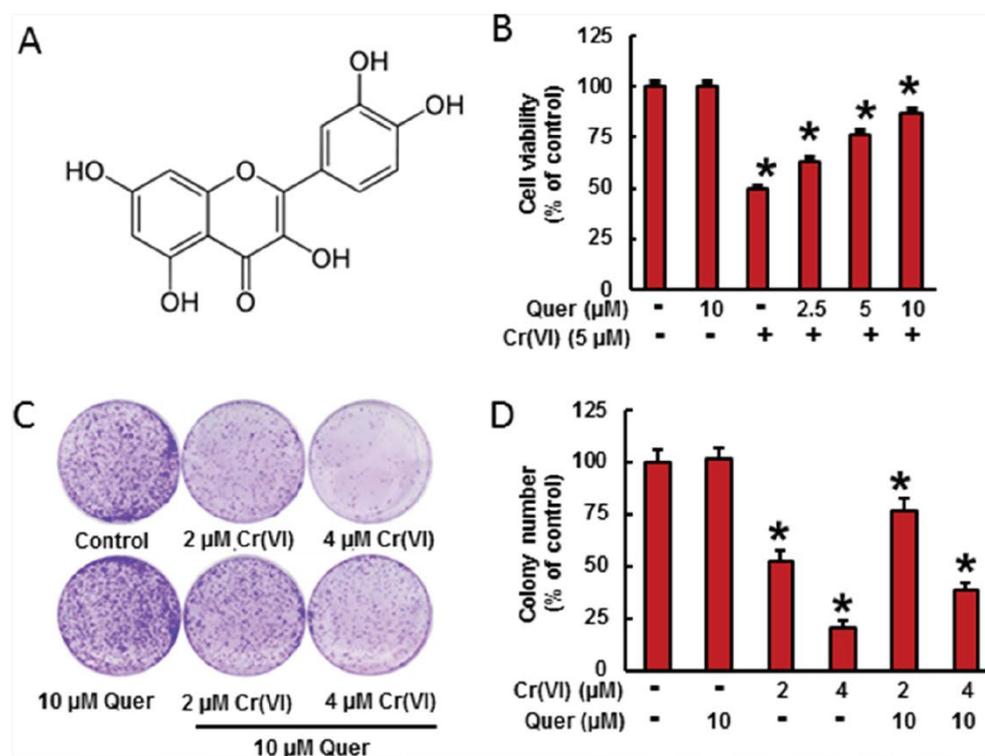


Figure 1: Quercetin inhibits Cr(VI)-induced cytotoxicity. A. Chemical structure of quercetin. B. BEAS-2B cells were treated with Cr(VI) (5 μM) for 24 h in the presence of quercetin (0, 2.5, 5, 10 μM). Cell viability was determined by MTT assay. C-D. BEAS-2B cells were treated with 2 μM or 4 μM Cr(VI) with or without 10 μM quercetin for 48 h, reseeded and cultured in drug free medium for an additional 7 days and stained with crystal violet. Colony numbers in the entire dish were counted. Data presented in the bar graphs are the mean ± SD of three independent experiments. *indicates a statistically significant difference from control cells with p<0.05.

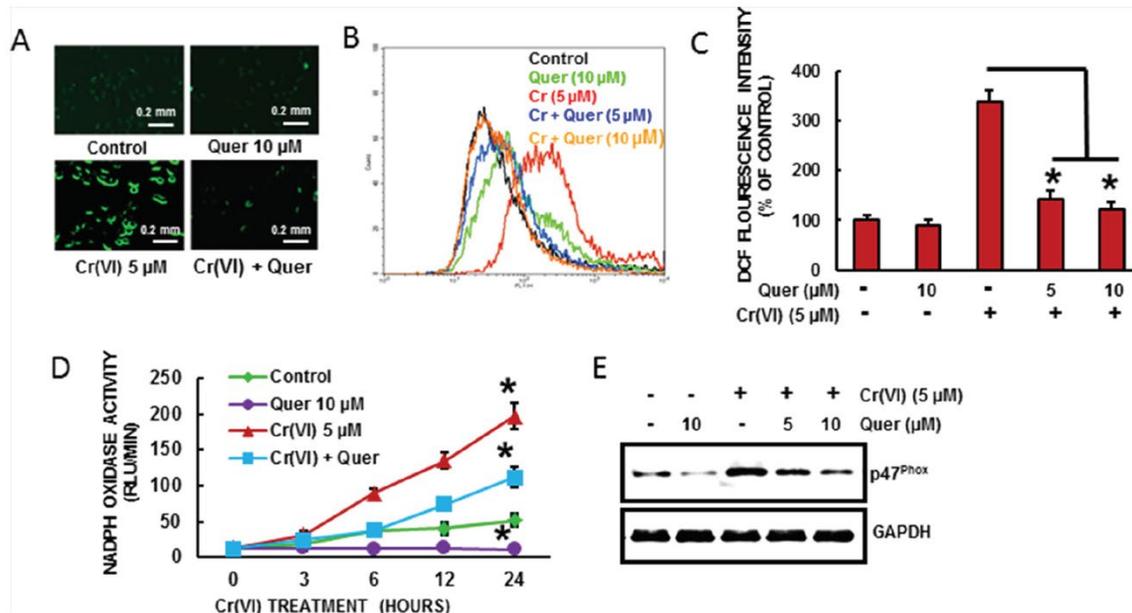


Figure 2: Quercetin inhibits Cr(VI)-induced ROS generation. BEAS-2B cells were exposed to Cr(VI) (0 or 5 μM) with or without quercetin (0, 5, 10 μM) for 12 h and then were labeled with A-C, DCFDA (10 μM). Images were taken with fluorescence microscopy and fluorescent intensity was determined by flow cytometry. **D.** NOX activity was measured by lucigenin chemiluminescence assay with Cr(VI) (0 or 5 μM) in the presence of quercetin (10 μM) for indicated times. **E.** Quercetin inhibits Cr(VI)-induced protein levels of NOX subunit, p47^{phox}. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *indicates a statistically significant difference from control cells with $p < 0.05$.

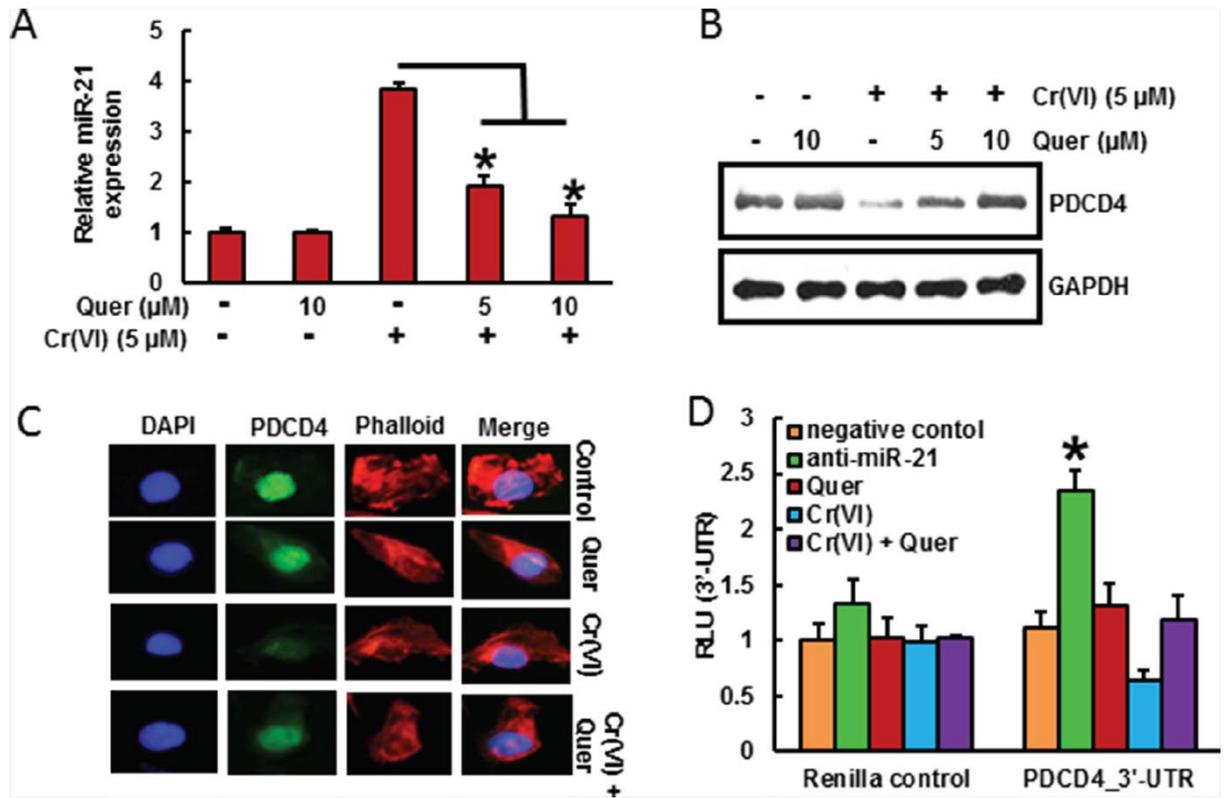


Figure 3: Quercetin inhibits Cr(VI)-induced miR-21 elevation and PDCD4 reduction. BEAS-2B cells were treated with Cr(VI) (5 μ M) for 24 h in the presence of quercetin (0, 5, 10 μ M). **A.** The relative miR-21 level was determined by Taqman real-time PCR. **B.** PDCD4 protein levels after acute Cr(VI) treatment was detected by immunoblotting. **C.** Representative images of fluorescence immunostaining of PDCD4. **D.** Quercetin ameliorates the Cr(VI)-induced inhibition of PDCD4 3'-UTR reporter activity. BEAS-2B cells were transfected with renilla reporter construct (pGL3-PDCD4_3'-UTR), miR-21 inhibitor (100 nM), negative control (100 nM), and pGL3-promoters and treated with 5 μ M Cr(VI) for 6 h in the presence of quercetin (10 μ M). Cellular lysates were subjected to a luciferase reporter assay as described in Materials and Methods. The results are expressed as relative activity (relative luminescence units (RLU)) normalized to the luciferase activity in the vector control cells without treatment. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *indicates a statistically significant difference from control cells with $p < 0.05$.

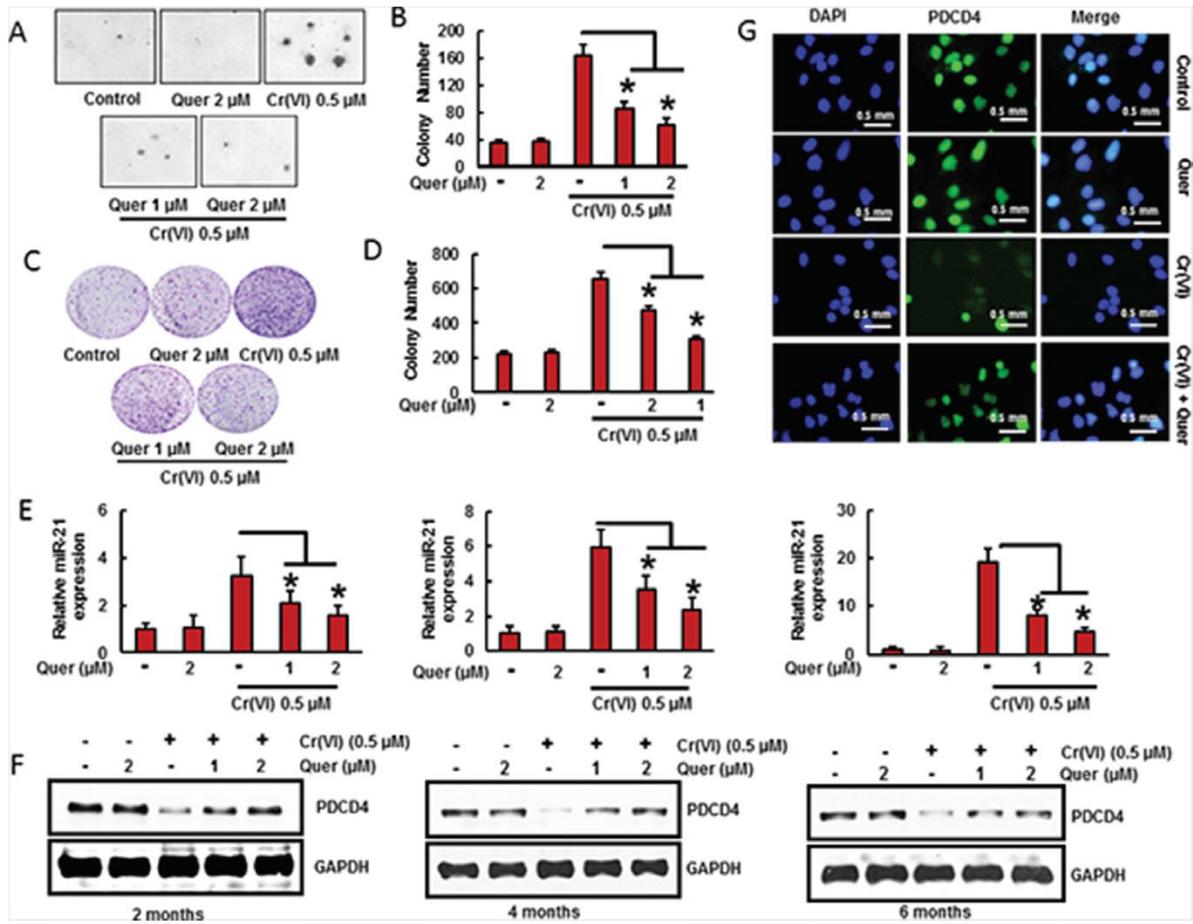


Figure 4: Quercetin inhibits chronic Cr(VI)-induced malignant cell transformation, miR-21 elevation and PDCD4 reduction. BEAS-2B cells were maintained in a medium containing Cr(VI) (0 or 0.5 μM) with or without quercetin (1 and 2 μM) for 6 months. **A-B.** Cells were cultured in 0.35% soft agar for 5 weeks. Colony numbers in the entire dish were counted. **C-D.** Cells cultured in drug free medium for an additional 7 days and stained with crystal violet. Colony numbers in the entire dish were counted. **E.** The relative miR-21 level was determined by Taqman real-time PCR. **F.** PDCD4 protein levels after co-treatment of quercetin with Cr(VI) was detected by immunoblotting **G.** Representative images of fluorescence immunostaining of PDCD4. Data presented in the bar graphs are the mean ± SD of three independent experiments. *indicates a statistically significant difference from control cells with p<0.05.

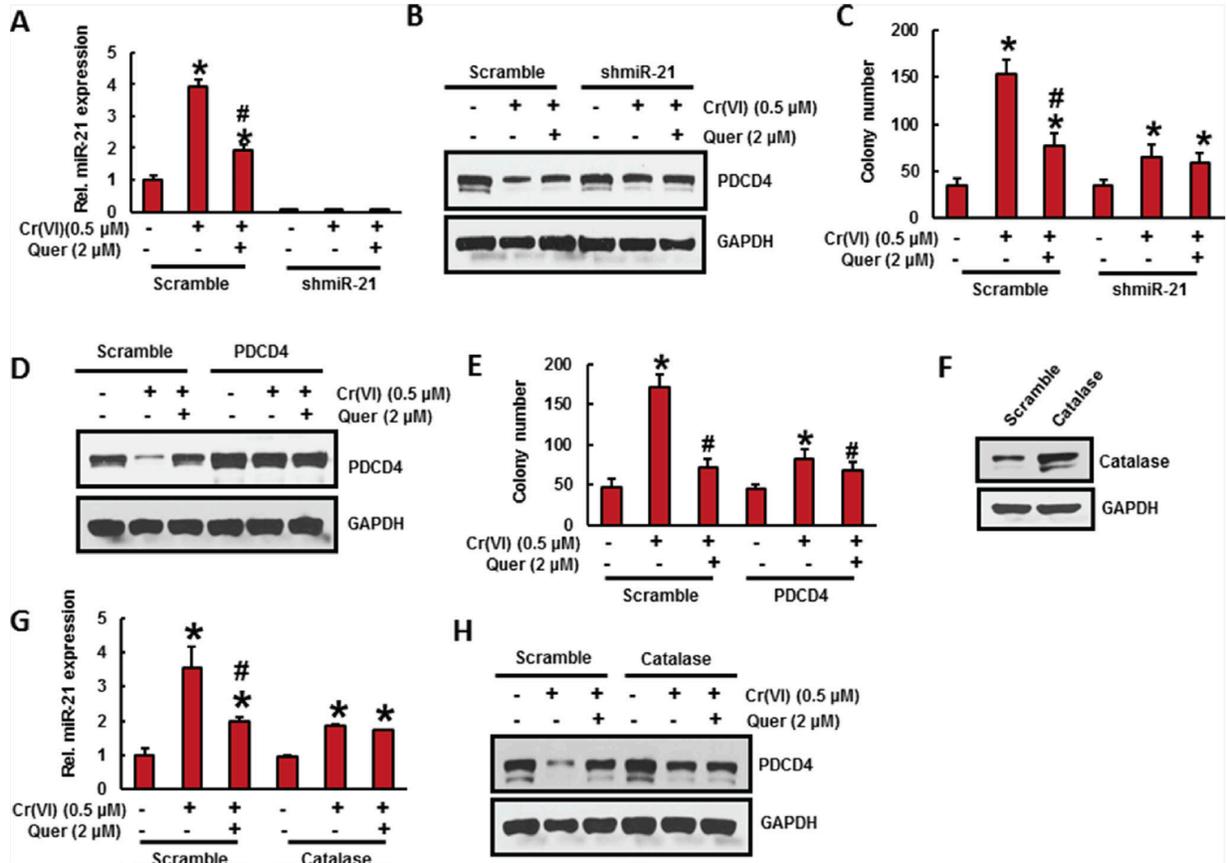


Figure 5: Stable knockdown of miR-21 and overexpression of PDCD4 or catalase in BEAS-2B cells significantly reduces the Cr(VI)-induced cell transformation. Stable knockdown of miR-21 in BEAS-2B cells suppresses the Cr(VI)-induced cell transformation. **A-C.** BEAS-2B cells were stably knockdown with miR-21 shRNA or their corresponding vehicle vector and treated with Cr(VI) (0 or 0.5 μ M) with or without quercetin (2 μ M) for 6 months. (A) The relative miR-21 level was determined by Taqman real-time PCR. (B) Cell lysates were prepared to determine the protein level of PDCD4 using Western blot analysis. (C) Malignant cell transformation was determined by soft agar assay. Stable overexpression of PDCD4 in BEAS-2B cells reduces the Cr(VI)-induced cell transformation. **D-E.** BEAS-2B cells were stably overexpressed with PDCD4 or their corresponding vehicle vector and treated with Cr(VI) (0 or 0.5 μ M) with or without quercetin (2 μ M) for 6 months. (D) Cell lysates were prepared to determine the protein level of PDCD4 using Western blot analysis. (E) Anchorage independent growth was determined by soft agar assay. Stable overexpression of catalase in BEAS-2B cells decreases the Cr(VI)-induced cell transformation. **F-H.** BEAS-2B cells were stably overexpressed with catalase or their corresponding vehicle vector treated with Cr(VI) (0 or 0.5 μ M) with or without quercetin (2 μ M) for 6 months. (F) BEAS-2B cells overexpressed with catalase was determined by Western blotting. (G) The relative miR-21 level was determined by Taqman real-time PCR. (H) PDCD4 protein level was detected by immunoblotting. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *# indicates a statistically significant difference from respective control cells with $p < 0.05$.

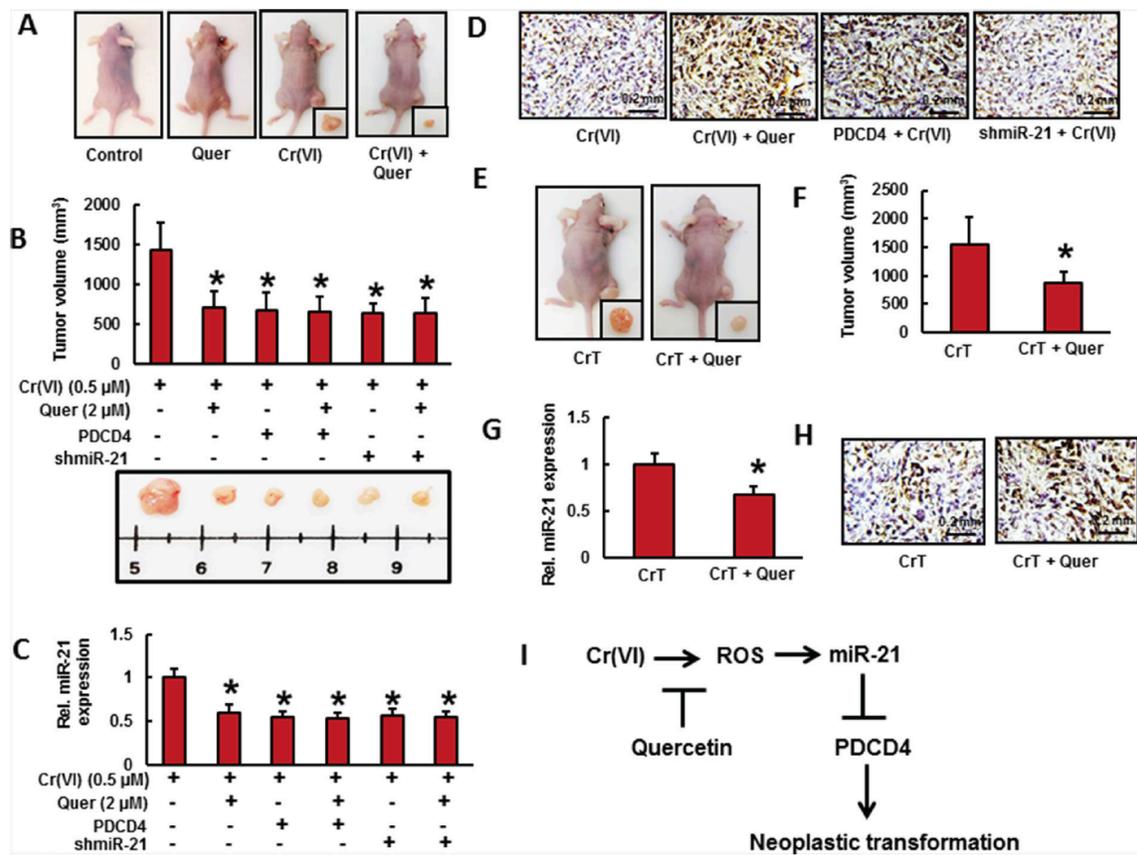


Figure 6: Quercetin inhibits the growth of xenograft tumors in mice from cells chronically exposed to Cr(VI). Cells from different treatments were injected into the flanks of 6-week old athymic nude mice (2×10^6 cells per mouse). Mice were checked daily for tumor appearance, and tumor volume was measured after 30 days. Tumor volume was determined by Vernier caliper, following the formula of $A \times B^2 \times 0.52$, where A is the longest diameter of tumor and B is the shortest diameter. **A.** Mice injected with BEAS-2B cells exposed to quercetin (2 μM) along with Cr(VI) (0.5 μM) showed reduced tumor incidence. **B.** Stable knockdown of miR-21 and overexpression of PDCD4 reduces the tumorigenicity of chronic Cr(VI) exposed BEAS-2B cells in nude mice. **C.** The relative miR-21 level was determined by Taqman real-time PCR. **D.** PDCD4 protein expression was detected by immunohistochemistry. **E.** CrT cells (2×10^6 per mouse) were injected into the 6-week-old female athymic nude mice. After the tumors had developed (about 100 mm³), the mice were injected with or without 10 mg/Kg/day quercetin (ip) every day for 30 days. **F.** Solid tumors in the quercetin treated mice were significantly smaller than those in the control mice. Consistent with tumor volume data, tumors from quercetin treated animals showed a decreased **G.** miR-21 level and more **H.** PDCD4 positive cells compared to tumors from control mice. Data presented in the bar graphs are the mean \pm SD of three independent experiments. * indicates a statistically significant difference from respective control with $p < 0.05$. **I.** Proposed mechanism of quercetin inhibits Cr(VI)-induced malignant cell transformation.

1.9c. *Data provided.* From the Respondents (for each figure):
 Data were provided for Figs. 1B&C (Excel spreadsheet data), 3D (Excel spreadsheet data), 5C&E (Excel spreadsheet data) (Appendix 100) but not for the other figures in the manuscript.

From the Respondents (Appendix 100): “All original data was generated by Dr. Poyil Pratheeshkumar. He was a postdoc from 2011-2016. He is now at Human Cancer Genomic Research, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. Dr. Pratheeshkumar provided the PPT data. Dr. Shi asked him for original data. His answer is that he stored the data in his home country, India, and cannot access it anytime soon.” – See email stating that in Appendix 100.

- 1.9d. *Our analysis.* Only a fraction of the original data was provided and the rest of it was unavailable. However, from the emails above, the original blots were left in the laboratory and could have been provided by the Respondents. The committee found, in the sequestered documents, some of Dr. Pratheeshkumar's lab notebooks and annotated western blotting films consistent with his email.
- 1.9e. *Relevant interview questions/comments.* Not applicable.
- 1.9f. *Conclusion for this specific example.* Based on the email exchange and the sequestered documents, the committee expected that more original data should have been provided by the Respondents. Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 1.10. M11: Roy, R.V., Pratheeshkumar, P., Son, Y.O., Wang, L., Hitron, J.A., Divya, S.P., Zhang, Z., Shi, X., Different roles of ROS and Nrf2 in Cr(VI)-induced inflammatory responses in normal and Cr(VI)-transformed cells. *Toxicology and Applied Pharmacology* 307, 81-90, 2016 (Appendix 023): Most figures
- 1.10a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 1.10b. *Original figures from Appendix 023 pages 83-88.*

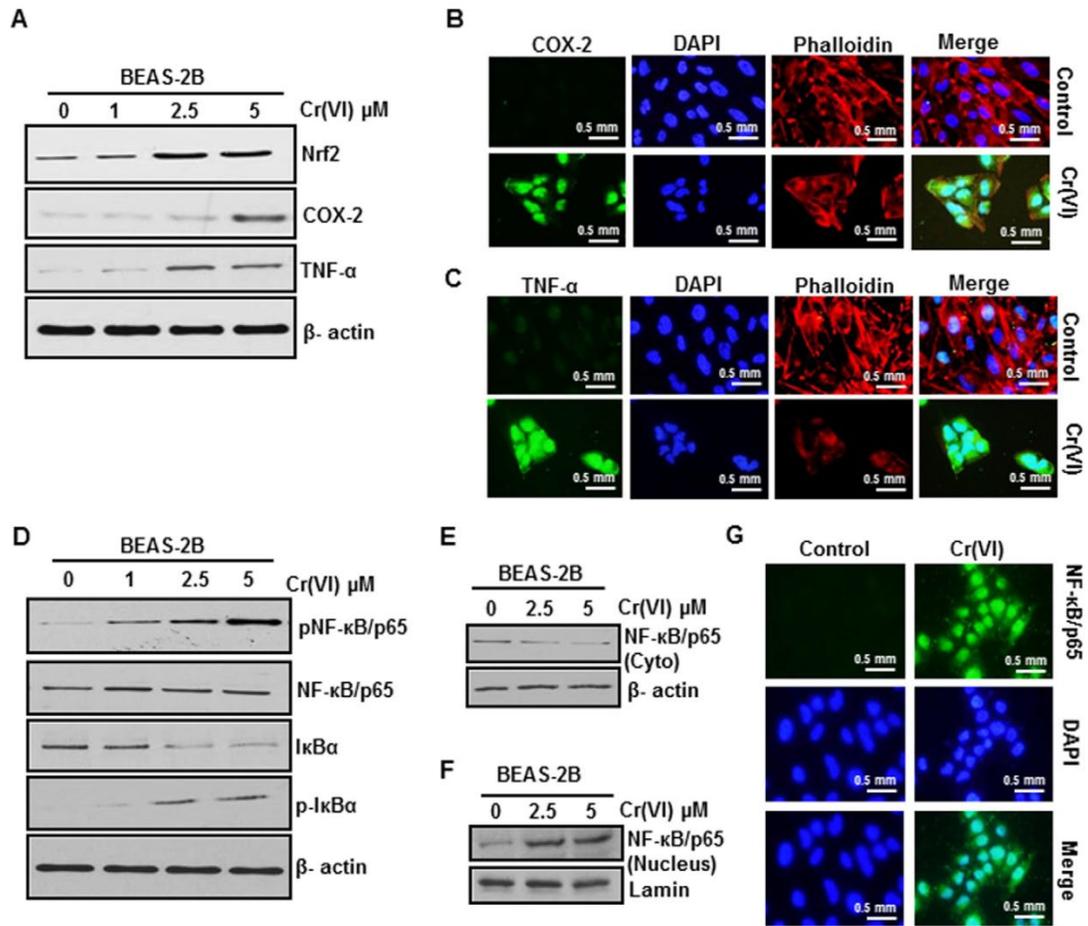


Fig. 1. Acute Cr(VI) stress induces inflammatory responses. (A) BEAS-2B cells were exposed to increasing concentrations of Cr(VI) (1, 2.5 and 5 μ M). After 24 h of exposure, total cells lysates were prepared and analyzed using Western blot to identify proteins associated with inflammatory responses (COX-2 and TNF- α) and antioxidant responsive gene markers (Nrf2). Representative Immunofluorescence images shows increased basal expression of (B) COX-2 and (C) TNF- α in Cr(VI)-treated BEAS-2B cells. (D) Total cell lysate analysis for NF- κ B signaling after Cr(VI) treatment. (E) cytoplasmic (NF- κ B-p65) and (F) Nuclear (NF- κ B-p65) analysis of NF- κ B signaling after Cr(VI) treatment. Lamin and β -actin were used as loading control for nuclear and cytoplasmic proteins. (G) Images of subcellular localization of NF- κ B-p65 under Cr(VI)-induced stress is shown. BEAS-2B cells were treated with Cr(VI) (5 μ M), and after 24 h were fixed and stained with NF- κ B-p65.

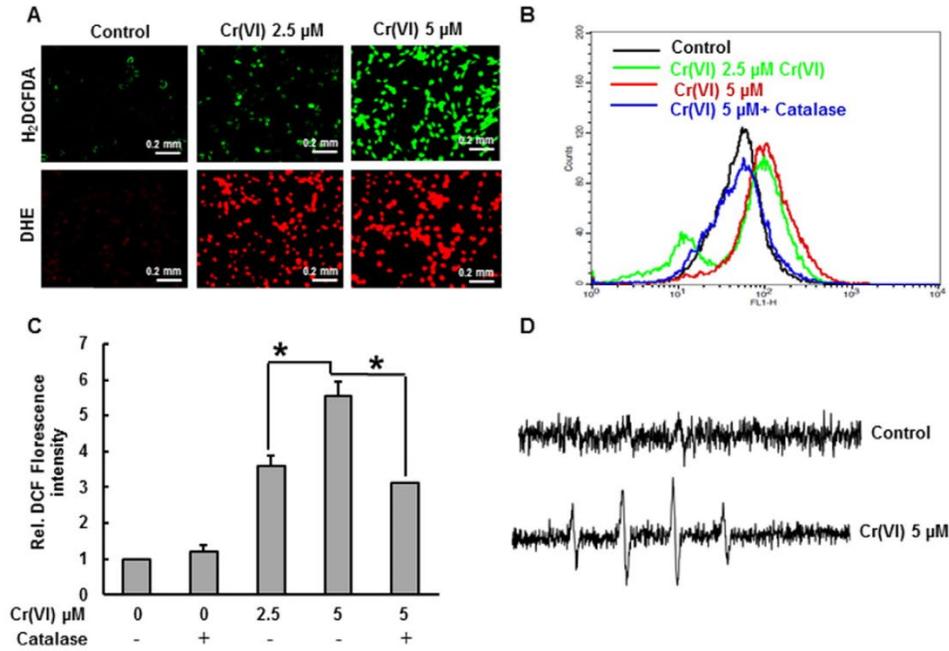


Fig. 2. Cr(VI) stress induces ROS generation in BEAS-2B cells. (A) BEAS-2B cells were exposed to Cr(VI) at indicated concentrations for 24 h and then stained with DCFDA or DHE (10 μ M), respectively for 30 min. Cells were imaged by fluorescence microscopy. (B–C) BEAS-2B cells or BEAS-2B cells overexpressing catalase were treated with indicated concentrations of Cr(VI) for 24 h and then exposed to DCFDA (10 μ M) and subjected to flow cytometer analysis and tabulated. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *indicates a statistically significant difference from control cells with $p < 0.05$. (D) Electron spin resonance (ESR) spectrum recorded 15 min after addition of 5 μ M of Cr(VI) to BEAS-2B cells to measure basal level of ROS. The generation of 1:2:2:1 quartet ESR signal is demonstrated.

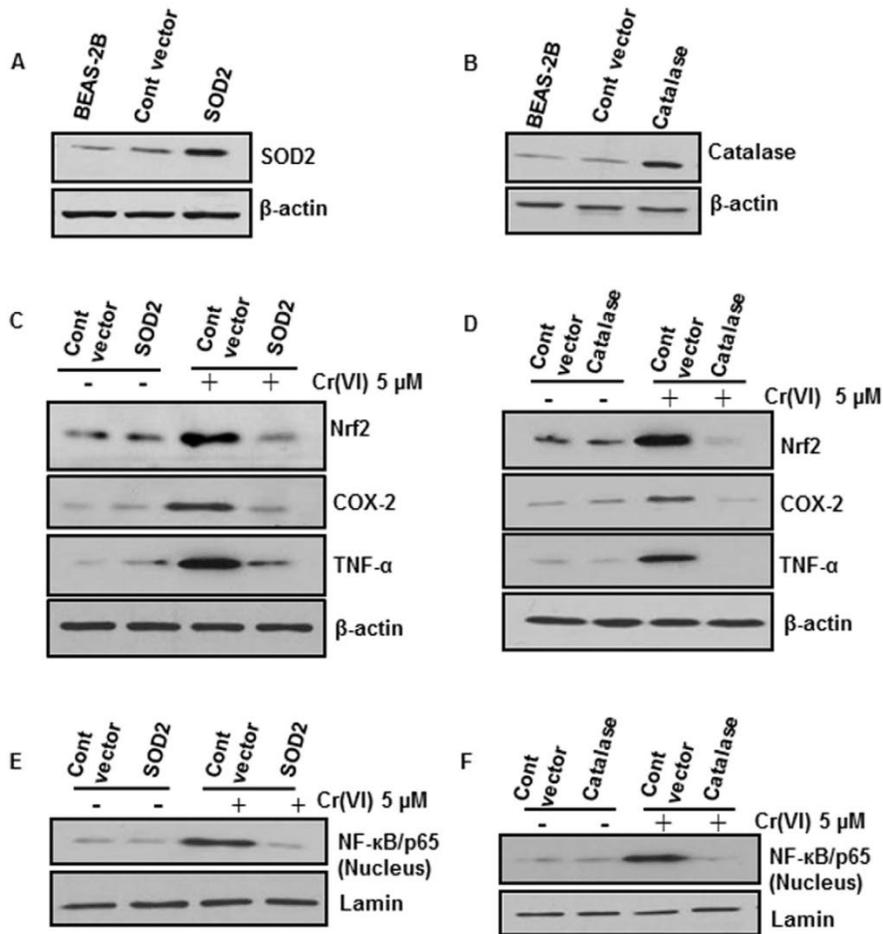


Fig. 3. ROS is responsible for acute Cr(VI)-induced inflammatory responses. Western Blot Analysis of BEAS-2B cells transiently transfected with (A) SOD2 or (B) CAT over-expressing plasmid. (C) BEAS-2B cells or BEAS-2B cells overexpressing SOD2 were treated with 5 μM of Cr(VI). After 24 h of exposure, total cells lysates were prepared and analyzed by Western blot to identify proteins associated with inflammatory responses (COX-2 and TNF-α) and Nrf2. (D) BEAS-2B cells or BEAS-2B cells overexpressing CAT were treated with 5 μM of Cr(VI) for 24 h and analyzed by Western blot to identify proteins associated with inflammation (COX-2 and TNF-α) and Nrf2. BEAS-2B cells or BEAS-2B cells overexpressing (E) SOD2 or (F) CAT were treated with 5 μM of Cr(VI) for 24 and analyzed for the nuclear translocation of NF-κB/p65.

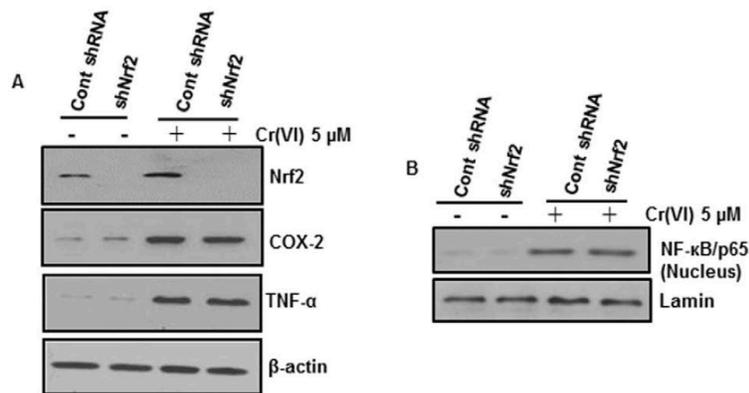


Fig. 4. Knock-down of Nrf2 has no effect on acute Cr(VI)-induced inflammatory responses. (A) BEAS-2B cells or Nrf2 silenced BEAS-2B cells were treated with 5 μM of Cr(VI). After 24 h of exposure, total cells lysates were prepared and analyzed by Western blot to identify proteins associated with inflammation (COX-2 and TNF-α) and Nrf2. BEAS-2B cells or (B) Nrf2 silenced BEAS-2B were treated with 5 μM of Cr(VI) for 24 and analyzed for the nuclear translocation of NF-κB/p65.

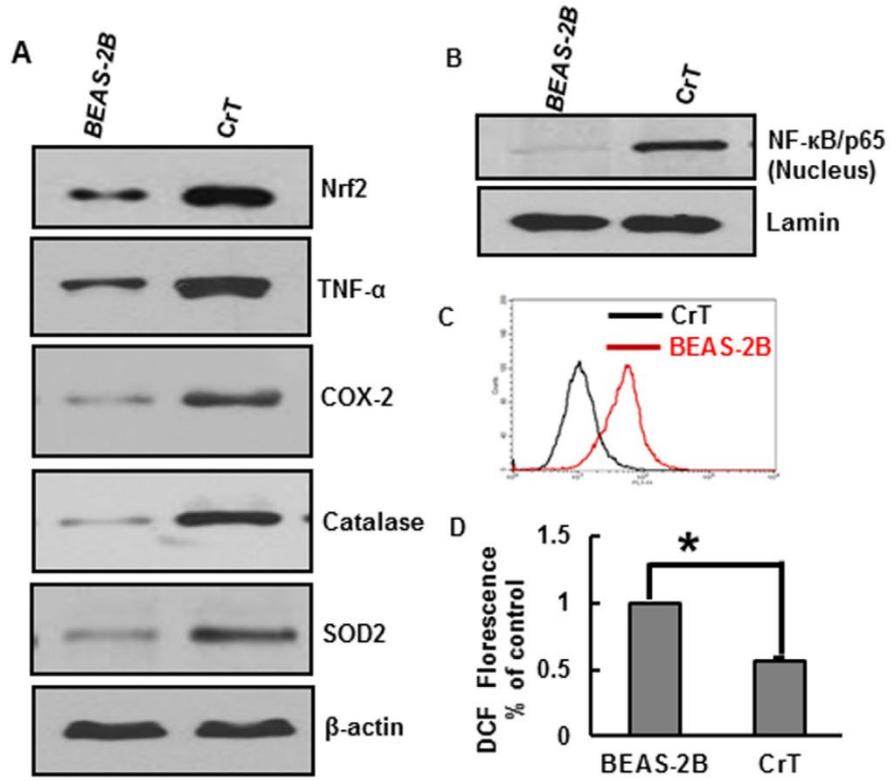


Fig. 5. High basal level expression of inflammatory proteins and low levels of ROS in Chromium transformed (CrT) cells. (A) Western blot analysis of basal levels of inflammatory (COX-2 and TNF- α) and antioxidant responsive gene markers (Nrf2) in Chromium transformed (CrT) cells and in passage matched BEAS-2B cells. (B) Nuclear accumulation of NF- κ B-p65 in BEAS-2B and CrT cells. (C-D) Quantitation of basal levels of ROS in BEAS-2B and CrT cells.

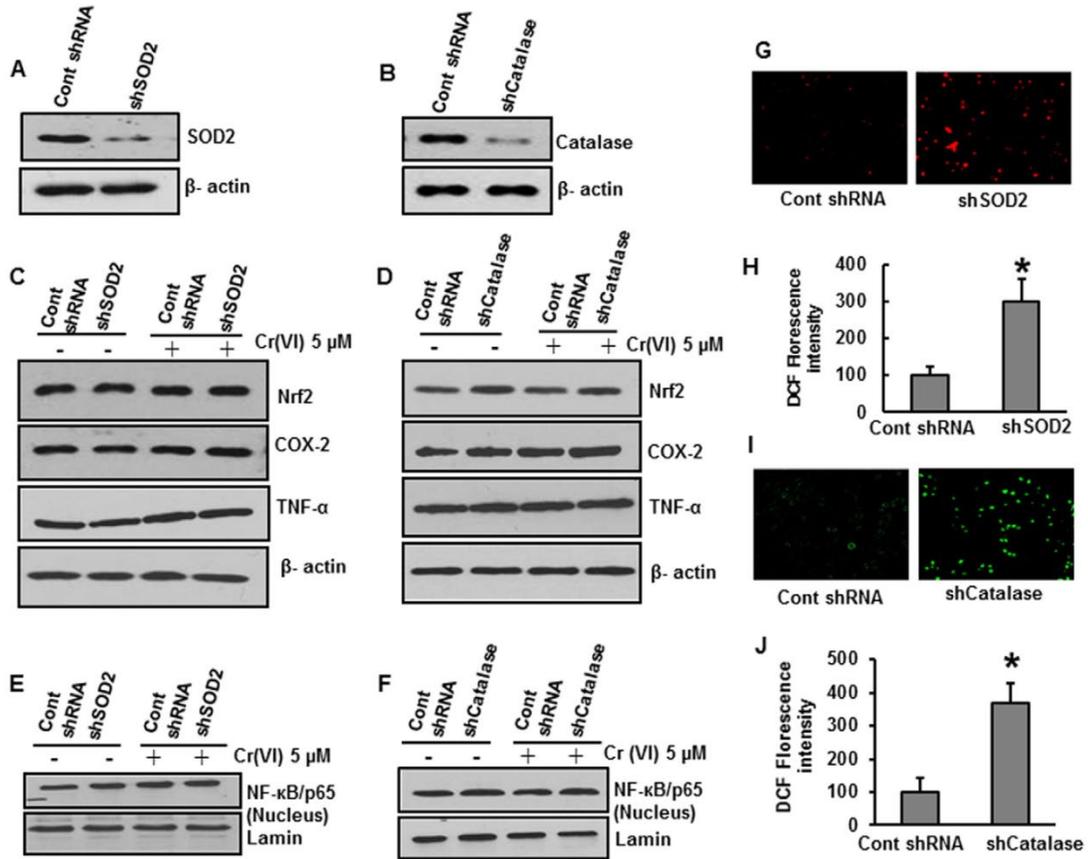


Fig. 6. ROS independent inflammation in CrT cells. Western blot analysis of CrT cells transiently transfected with (A) shSOD2 and (B) shCatalase plasmid. CrT cells silenced with (C) shSOD2 and (D) shCatalase plasmid were treated with or without Cr(VI) for 24 h, total cell lysates were prepared and analyzed using Western blot for inflammatory markers (COX-2 and TNF- α) and Nrf2. CrT cells or (E) SOD2 silenced CrT or (F) Catalase silenced CrT were treated with 5 μ M of Cr(VI) for 24 and analyzed for the nuclear translocation of NF- κ B/p65. CrT cells were transfected with (G) shSOD2 or (I) shCatalase and analyzed for ROS level. Cells were labeled with (G-H) DHE (10 μ M) or (I-J) DCFDA (10 μ M). Images were obtained by fluorescence microscopy and fluorescent intensity was determined by flow cytometry. Data presented in the bar graphs are the mean \pm SD of three independent experiments. * indicates a statistically significant difference from control cells with $p < 0.05$.

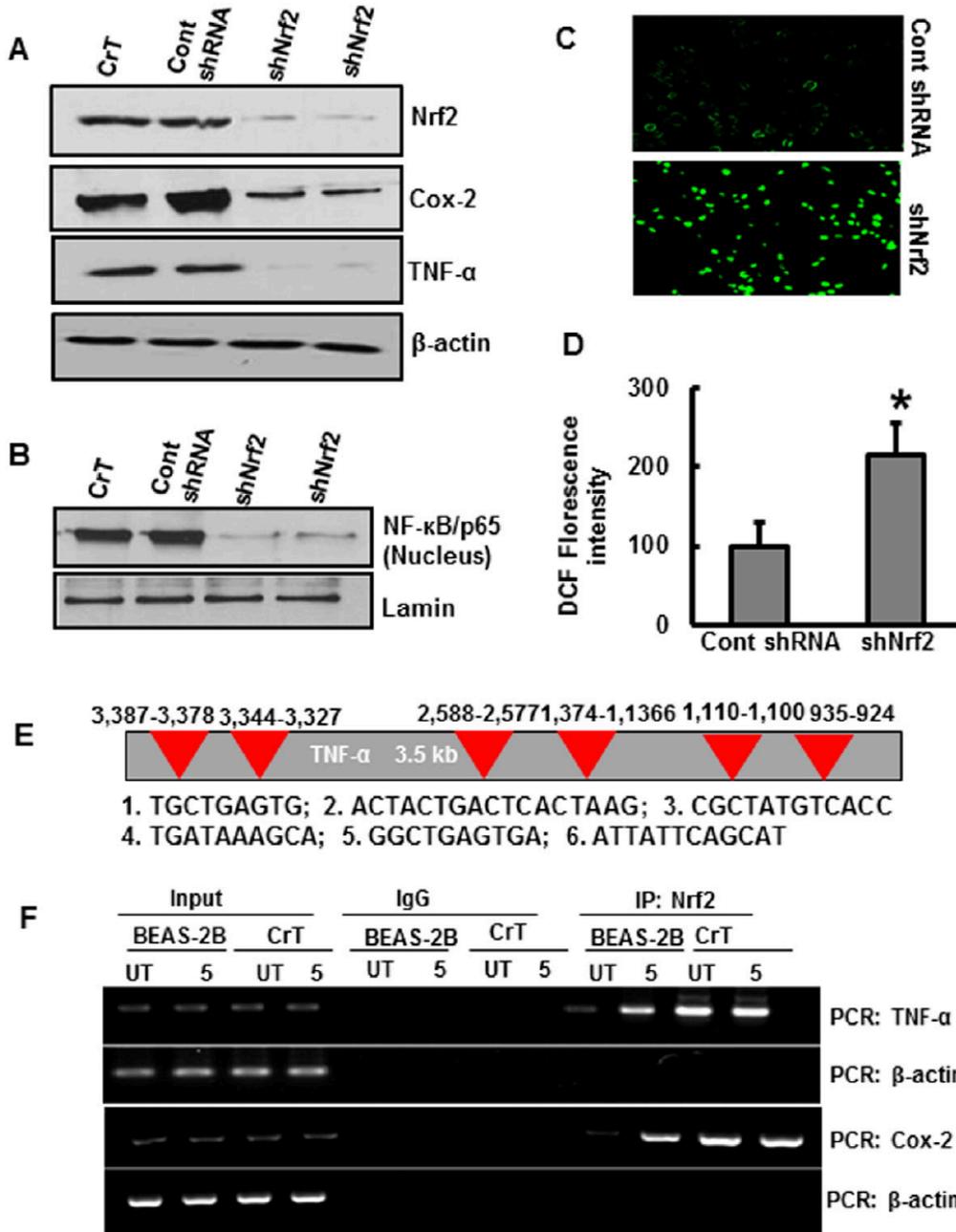


Fig. 7. Constitutively high levels of Nrf2 regulate inflammation. (A) CrT cells were transfected with shNrf2 and total cells lysates were prepared and analyzed for inflammatory markers (COX-2 and TNF- α). (B) Nuclear accumulation of NF- κ B-p65 in CrT cells after knock-down with shNrf2. (C) CrT cells were transfected with shNrf2 and basal ROS level were measured. Cells were labeled with DCFDA (10 μ M). Images were obtained by (C) fluorescence microscopy and fluorescent intensity was determined by (D) flow cytometry. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *indicates a statistically significant difference from control cells with $p < 0.05$. (E). Consensus or putative ARE regions of TNF- α promoter in the 3.5 kb region. (F) ChIP analysis for increased basal Nrf2 associating with TNF- α and COX-2 in CrT cells were carried out by precipitating genomic DNA with Nrf2.

1.10c. *Data provided.* Some of the panel images for Figs. 1B, C and G were provided as JPG files. Some data for Figs. 2B (as FACS printout), 3C,D (scans of full gels), 4A,B (scans of full gels), 5C (as FACS printout), 6C,D (scans of full gels), and 7F (full agarose gels) (Appendix 101). The data provided were in an electronic format only.

1.10d. *Our analysis.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092).

1.10e. *Relevant interview questions/comments.* Not applicable.

1.10f. *Conclusion for this specific example.* Data were not retained based on UK (Appendix 091) and federal guidelines (Appendix 092). This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

2. Inappropriate loading controls:

Description: The committee defined this category as figures presented with incorrect loading controls, controls done months apart from the experiments, and/or the same controls used for multiple different experiments.

Overall summary: 4 figures from grants and 6 figure panels from manuscripts (sections 2.1-2.8).

Overall summary of interviews and relevant comments: During the interviews, the committee asked each Respondent, as well as two other laboratory personnel (Drs. Hitron and Wang), a series of questions about which loading controls were valid and how they were used. There was some variability regarding the load controls used: GAPDH, β -actin and α -actinin (Appendices 049 (Kim interview I, page 54); 052 (Shi interview, page 53); 054 (Zhang interview, pages 56-57); and 051 (Hitron interview, page 45). Dr. Zhang stated that she “requires her employees to have a loading control” (Appendix 054 (Zhang interview, pages 70-71)). Dr. Shi stated that “he wants to see loading controls in each gel” (Appendix 052 (Shi interview, pages 79-80)).

Overall conclusions: Despite the Respondents’ claims that they request loading controls to be performed for each gel, the committee observed numerous instances where the same loading control was inappropriately used for multiple figures. In some cases, loading control blots were performed months apart from the other blots shown in the figure. In addition, the committee identified cases where loading controls were inappropriately labeled (α -actinin was shown but labeled as β -actin). Based on these findings and the withdrawal notes associated with three retracted manuscripts (Appendices 055, 056, and 057) and PubPeer reports (Appendix 138), the misuse of loading controls is a recurring and systemic problem in the Respondents’ laboratory.

Specific examples for “inappropriate loading controls”:

2.1. G2: Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Figs. 9 and 10

2.1a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

2.1b. *Original figure from Appendix 007 page 4 of the research strategy section of the grant.*



Fig. 9 Luteolin increases Nrf2 expression in normal cells but decreases constitutive Nrf2 expression in Cr(VI)-transformed cells. Normal parent BEAS-2B cells and Cr(VI)-transformed BEAS-2B cells (B2B-Cr) were treated with luteolin (10 μ M) for 24 hours. Cells were harvested and whole protein lysates were extracted. Nrf2 expression was examined by immunoblotting.

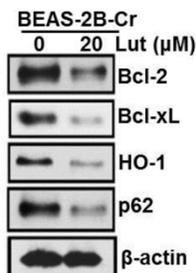


Fig. 10 Luteolin inhibits constitutive expressions of Bcl-2, Bcl-XL, HO-1, and p62 in Cr(VI)-transformed cells. Cr(VI)-transformed BEAS-2B cells (BEAS-2B-Cr) cells were treated with luteolin (20 μ M) for 24 hours. Cells were harvested and whole protein lysates were extracted. Expressions of Bcl-2, Bcl-XL, HO-1, and p62 were examined by immunoblotting.

- 2.1c. *Data provided.* No data were provided (sections 1.2 and 1.3 and corresponding Appendices 093 and 094).
- 2.1d. *Our analysis.* Fig. 9 and Fig. 10 are represented as two distinct experiments with two different concentrations of luteolin (10 μ M in Fig. 9, 20 μ M in Fig. 10). These luteolin concentrations were indicated in the figure labels and the associated figure legends. Based on the shape of the bands, background of the images, and the “fuzziness” around the bands, the committee concludes that the β -actin controls in Figs. 9 and 10 are the same images.

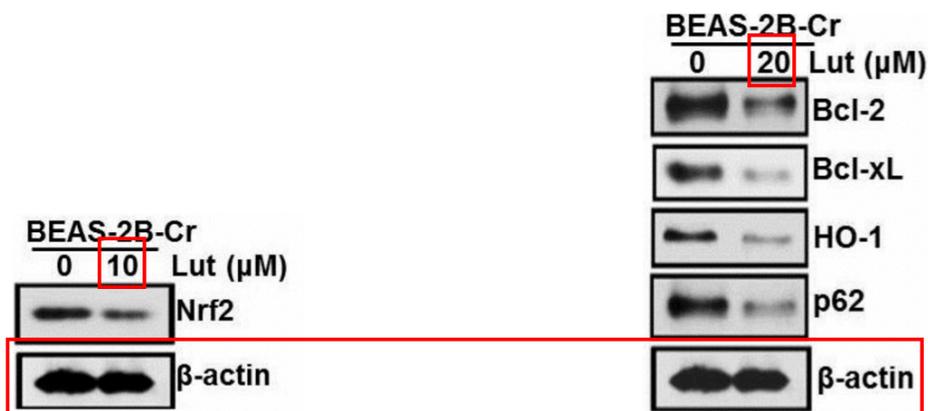


Fig. 9 Luteolin increases Nrf2 expression in normal cells but decreases constitutive Nrf2 expression in Cr(VI)-transformed cells. Normal parent BEAS-2B cells and Cr(VI)-transformed BEAS-2B cells (B2B-Cr) were treated with **luteolin (10 μ M)** for 24 hours. Cells were harvested and whole protein lysates were extracted. Nrf2 expression was examined by immunoblotting.

Fig. 10 Luteolin inhibits constitutive expressions of Bcl-2, Bcl-XL, HO-1, and p62 in Cr(VI)-transformed cells. Cr(VI)-transformed BEAS-2B cells (BEAS-2B-Cr) cells were treated with **luteolin (20 μ M)** for 24 hours. Cells were harvested and whole protein lysates were extracted. Expressions of Bcl-2, Bcl-XL, HO-1, and p62 were examined by immunoblotting.

- 2.1e. *Relevant interview questions/comments.* During the interview, Dr. Shi stated that the loading controls were provided by Dr. Pratheeshkumar (Appendix 052 (Shi interview, page 55)). Dr. Shi agreed that the loading controls for these two figures looked the same (Appendix 052 (Shi interview, page 55)). However, Dr. Shi also stated that there was a mistake in the

scale (*i.e.*, the label) in Fig. 9, which should be 20 μ M (Appendix 052 (Shi interview, page 55)). Dr. Shi believes that the same samples were used for both figures (Appendix 052 (Shi interview, page 56)). Dr. Shi further stated that he does not believe that this mislabeling changes the conclusions from the two experiments (Appendix 052 (Shi interview, page 56)).

During the interview, Dr. Shi brought in Exhibits L and M (Appendix 53), indicating that they were relevant to Figs. 9 and 10. Dr. Shi used these documents to demonstrate that the concentration of luteolin in Fig. 9 should be 20 μ M (Appendix 052 (Shi interview, pages 57-59)).

2.1f. *Conclusion for this specific example.* While likely a mistake in figure labeling, which does not change the conclusion of the experiment, this is an example of a recurrent lack of rigor in figure construction. This is a departure from accepted practices of the research community, but the committee could not determine whether research misconduct has occurred.

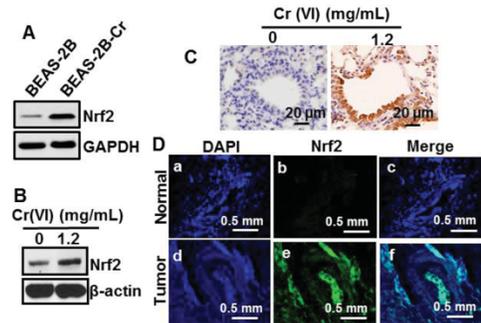
2.2. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 7A

2.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

2.2b. *Original figure from Appendix 008 page 54.*

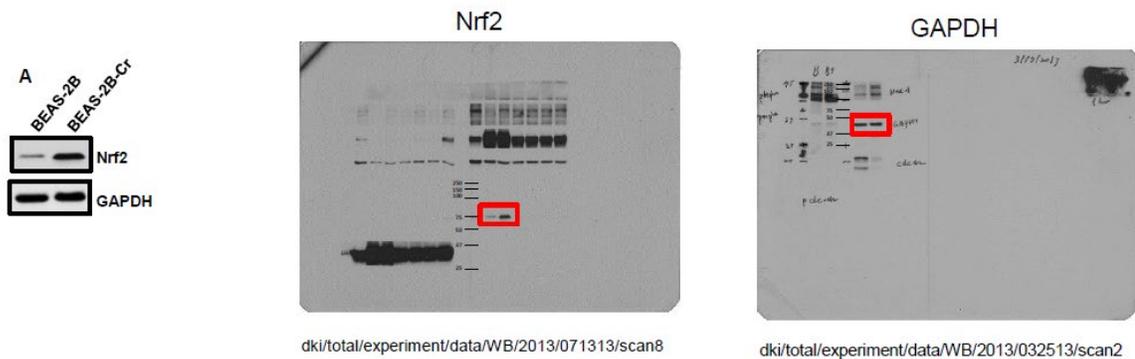
Fig. 7 Increased Nrf2 protein level in Cr(VI)-transformed cells, in lung tissues from Cr(VI)-exposed animals, and in lung tumor tissues from a worker exposed to Cr(VI) for 19 years. (A) Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) were cultured in 10-cm dishes. Whole protein was isolated for examination of Nrf2 level using immunoblotting. The results are representative of three independent experiments. (B) and (C) 4-week old both male and female BALB/c mice were intranasally exposed to Cr(VI) particles (1.2 mg/mL) as described in Fig. 2. Lung tissues were isolated to examine Nrf2 level using immunoblotting (B) and immunohistological (C) analyses. Intensity of brown color represents Nrf2 level (C). Sections are representative of 6 different animals from each treatment group. (D) Formalin-fixed human lung tumor and adjacent normal tissues from a non-smoking worker exposed to Cr(VI) were subjected to fluorescence immunostainings with DAPI (nuclear control, a and d) and Nrf2 (b and e). The merged images are provided in c and f. The results are representative of three independent experiments.



2.2c. *Data provided.* See Appendix 102.

From the Respondents:

Fig.7



- 2.2d. *Our analysis.* According to the file paths provided by the Respondents (file paths under the film scans in the figure above), the Nrf2 blot in this figure was imaged on 07/13/13, while the GAPDH load control was imaged on 03/25/13. Hand-written on the GAPDH gel is the date of 3/17/13. No such hand-written annotation was provided on the Nrf2 image.
- 2.2e. *Relevant interview questions/comments.* Not applicable.
- 2.2f. *Conclusion for this specific example.* The GAPDH loading control blot was performed months apart from the other blot shown in the figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.
- 2.3. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 8B
- 2.3a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 2.3b. *Original figure from Appendix 008 page 54.*

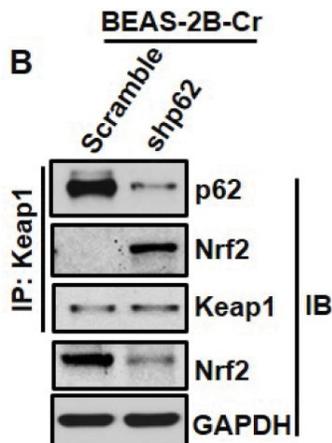
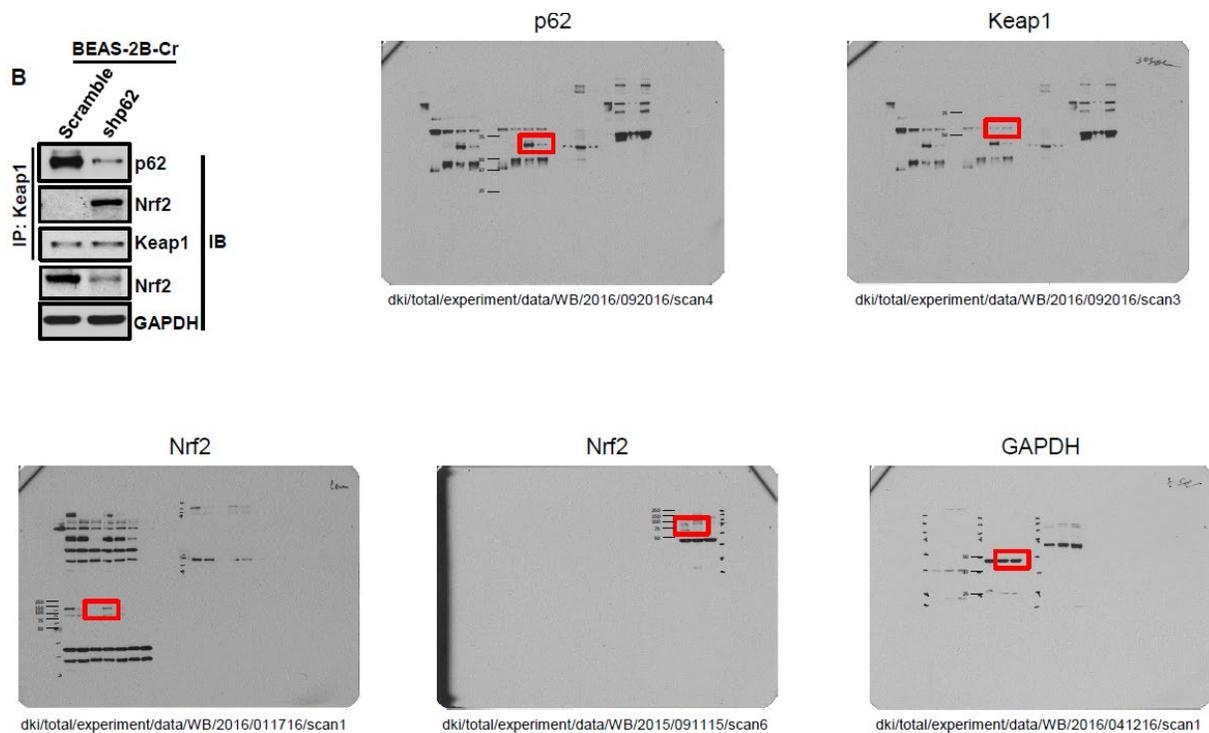


Fig. 8 p62 upregulates Nrf2 in Cr(VI)-transformed cells. (A) and (B) Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) with (shp62) and without (Scramble) knockdown of p62 by its shRNA were subjected to either co-immunoprecipitation and immunoblotting analyses. The results are representative of three independent experiments.

- 2.3c. *Data provided.* See Appendix 100.
 From the Respondents:



2.3d. *Our analysis.* According to the file paths provided by the Respondents (file paths under the film scans in the figure above), the images used for this figure were scanned on: p62 - 9/20/16, Keap1 - 9/20/16, 1st Nrf2 - 1/17/16, 2nd Nrf2 - 09/11/15 and GAPDH - 04/12/16. There was no relevant hand-written annotation.

2.3e. *Relevant interview questions/comments.* Not applicable.

2.3f. *Conclusion for this specific example.* The GAPDH loading control blot was performed months apart from the other blot shown in the figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

2.4. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 2C

2.4a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

2.4b. *Original figure from Appendix 013, page 394.*

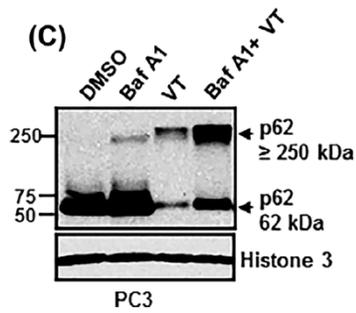
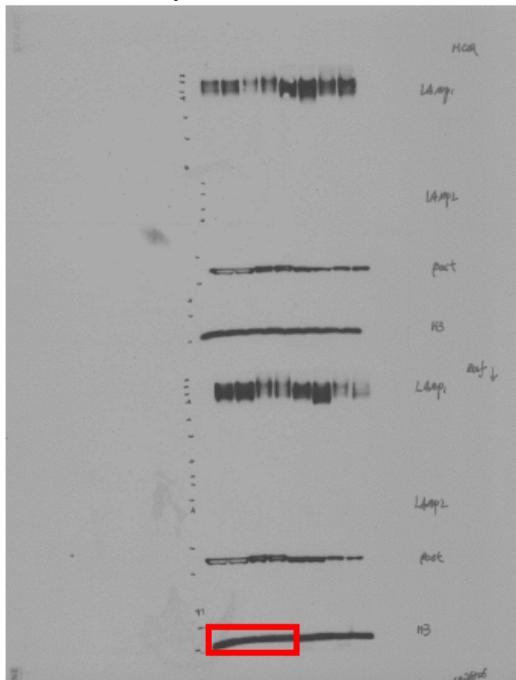


FIGURE 2 Verteporfin causes p62 crosslink in contrast to hydroxychloroquine and bafilomycin A1 do not. (A) and (B) PC-3 and LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), or their combination for 24 h. (C) and (D) PC-3 and LNCaP cells were treated with 10 nM bafilomycin A1 (Baf A1), 10 μ M Verteporfin (VT), or their combination for 24 h. The whole-cell lysates were collected for examination of p62 level using immunoblotting analysis in PC-3 (A and C) and LNCaP (B and D) cells. Histone 3 was used as a loading control. The results represent three independent experiments

2.4c. *Data provided.* See Appendix 104.
From the Respondents:



File name: loading

2.4d. *Our analysis.* Based on the shape, curvature and segmentation of the bands, the data provided for Histone 3 do not match the Histone 3 load control used in the figure (below).

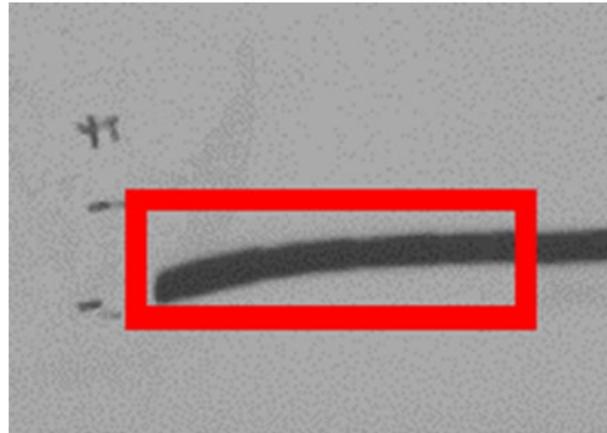
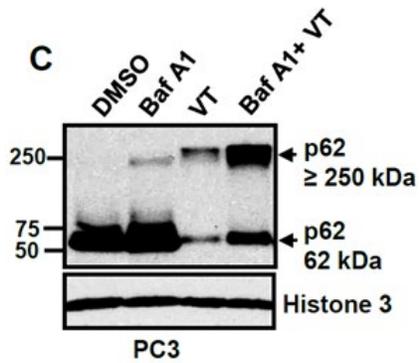


Fig 2

2.4e. *Relevant interview questions/comments.* Drs. Shi, Zhang and Wang were all asked about this figure and the data therein. They all agreed that Dr. Wang did the experiment represented (Appendices 052 (Shi interview, page 35); 054 (Zhang interview, page 32); and 058 (Wang interview, page 25)). Dr. Shi stated that the data provided in response to the committee’s request were not the correct exposures (Appendix 052 (Shi interview, pages 68-70)). Dr. Zhang also stated that the data did not match (Appendix 054 (Zhang interview, pages 62-63)).

Dr. DiPaola confirmed that the loading controls in the Figure did not match the data provided (Appendix 060 (DiPaola interview II, page 15)) and also stated that when Dr. Wang showed him the original Figure, it looked “convincing” (Appendix 060 (DiPaola interview II, page 16)).

Dr. Wang’s Exhibit D to his interview (Appendix 059), which was his lab notebook and contained relevant information, was not in the sequestered materials originally provided by the Respondents at the beginning of the investigation. Dr. Zhang stated that she was unable to find this data (Appendix 054 (Zhang interview, page 63)).

In Exhibits E-J to his interview (Appendix 059), Dr. Wang demonstrated how the original data sent to the committee was a mistake (Appendix 058 (Wang interview, pages 43-45)) and then tried to explain how the load controls were generated for Figure 2C (Appendix 058 (Wang interview, pages 47-48)). He started by indicating that the original data provided had the wrong bands boxed in red. Then Wang showed, using the exhibits to his interview, where the correct bands were on the gels and how he used them to create the Figure. Exhibit J to Wang’s interview (Appendix 059) was a step by step narrative of how the Figure was generated and what manipulations were done to the images. Dr. Wang stated that he now had the lanes correct (Appendix 058 (Wang interview, pages 50-52)). After all this explanation, the bands indicated in Exhibit J, panel D, do not match what was in the published Figure 2C. Moreover, in Exhibit J (Appendix 059), it was unclear how Dr. Wang went from Step C to Step D.

2.4f. *Conclusion for this specific example.* The data provided for Histone 3 do not match what was used in the figure, and the interviews and exhibits provided did not clarify how the published load control was generated. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation

is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 2.5. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013), Fig. 2D
- 2.5a. October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 2.5b. *Original figure from Appendix 013, page 394.*

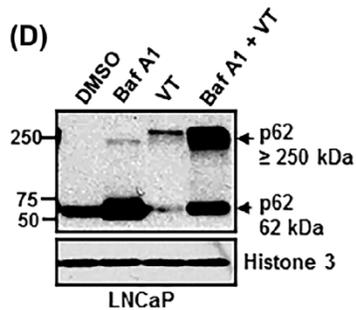
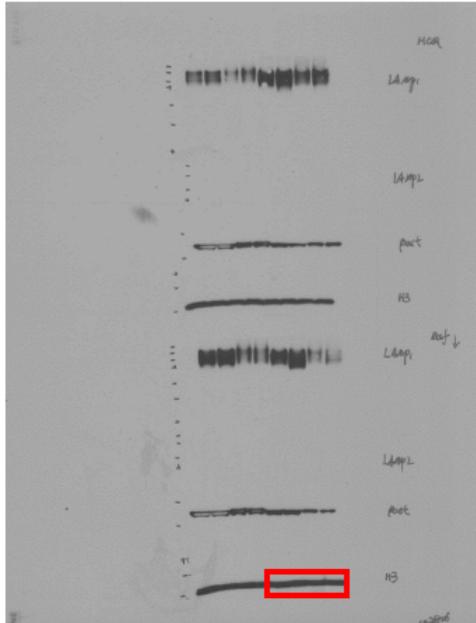


FIGURE 2 Verteporfin causes p62 crosslink in contrast to hydroxychloroquine and bafilomycin A1 do not. (A) and (B) PC-3 and LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), or their combination for 24 h. (C) and (D) PC-3 and LNCaP cells were treated with 10 nM bafilomycin A1 (Baf A1), 10 μ M Verteporfin (VT), or their combination for 24 h. The whole-cell lysates were collected for examination of p62 level using immunoblotting analysis in PC-3 (A and C) and LNCaP (B and D) cells. Histone 3 was used as a loading control. The results represent three independent experiments

2.5c. *Data provided.* See Appendix 105
From the Respondents:



File name: loading

2.5d. *Our analysis.* Based on the shape, curvature, segmentation of the bands and fuzziness around the bands, the data provided for Histone 3 do not match the Histone 3 load control used in the figure.

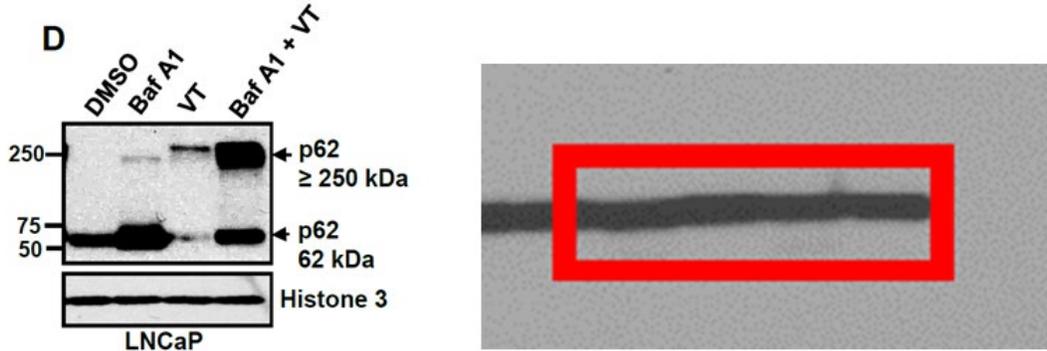


Fig 2

2.5e. *Relevant interview questions/comments.* Not applicable.

2.5f. *Conclusion for this specific example.* The data provided for Histone 3 do not match what was used in the figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

2.6. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013), Fig. 3D

2.6a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

2.6b. *Original figure from Appendix 013, page 395.*

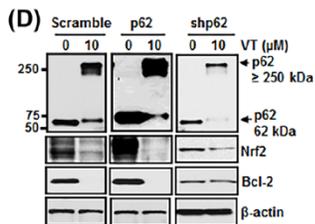
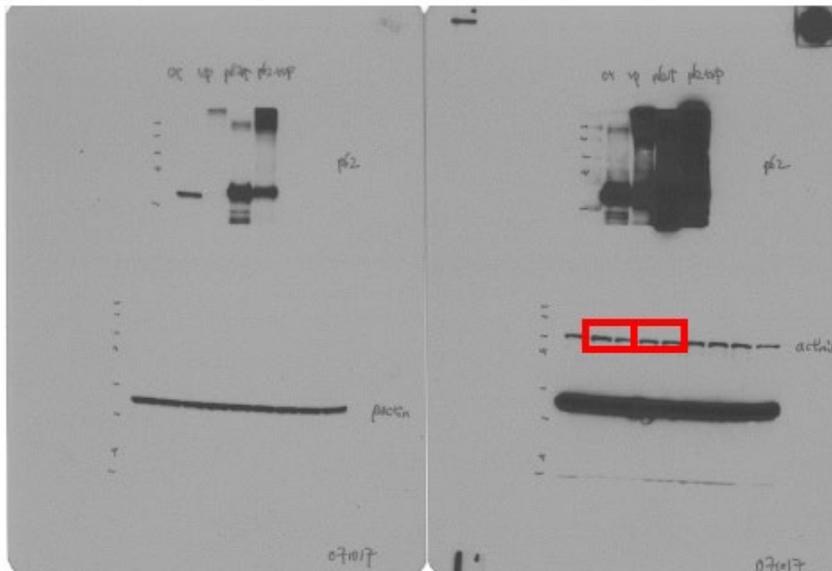


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μ M Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μ M and 10 μ M Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μ M 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean \pm SE ($n = 6$). * and #, $P < 0.05$ compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μ M of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

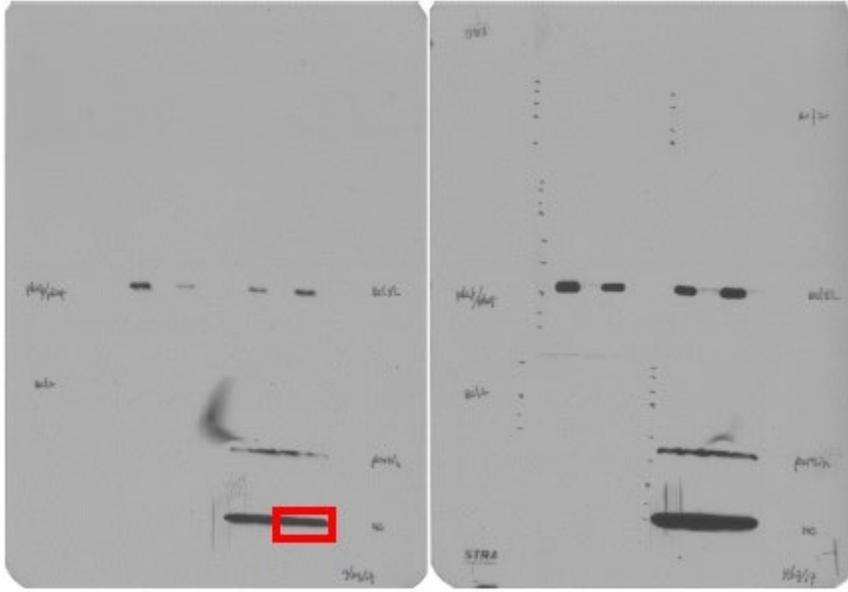
2.6c. *Data provided.* See Appendix 103.

Scan of full gel showing the loading control for the scramble and p62 overexpression conditions provided by the Respondents:



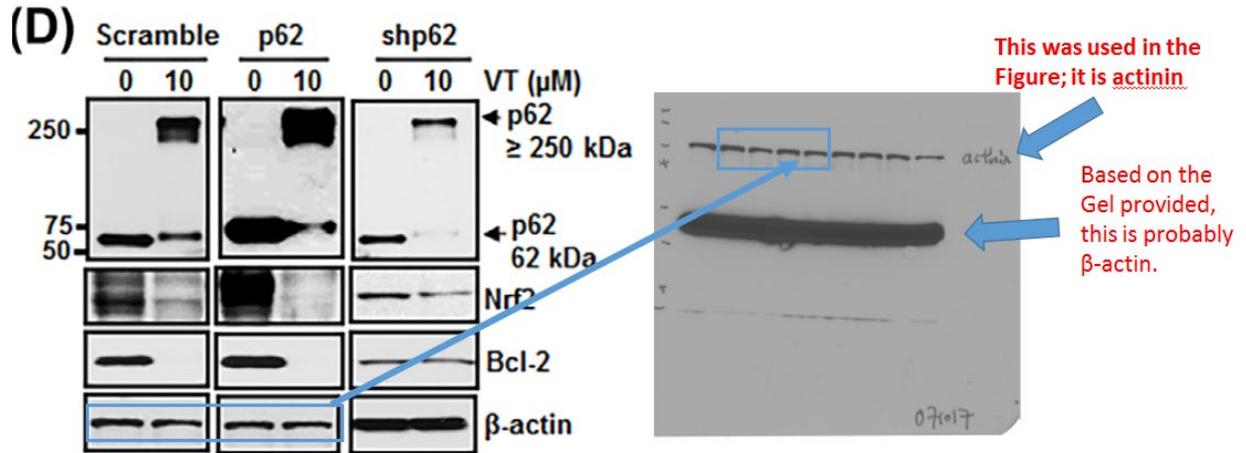
File name: loading 1

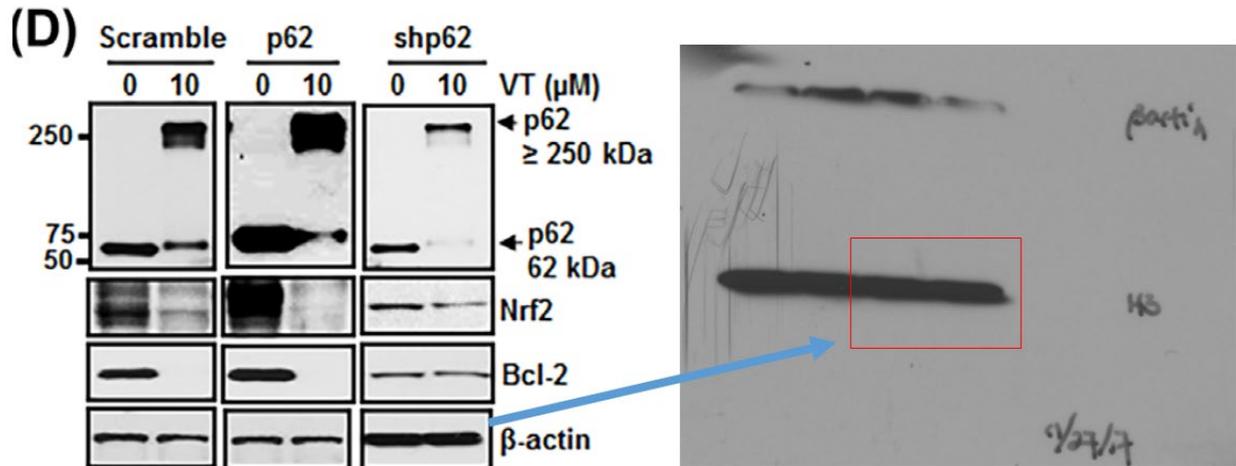
Scan of full gel showing the loading control for the shp62 condition provided by the Respondents:



File name: loading 2

2.6d. *Our analysis.* The hand-written annotations for the western blots were as follows: Nrf2 – 11/13/17, Histone 3 – 11/27/17, actinin – 07/10/17, p62 – 11/13/17, Nrf2 for scramble and p62 overexpression – 7/10/17, p62 for p62 overexpression - 7/14/17, Bcl2 for scramble and p62 overexpression – 07/13/17. Based on the data provided by the Respondents, the load controls used for SCRAMBLED and p62 conditions were actinin, but they are called β -actin in the figure (top panel below). The Respondents boxed Histone 3 as the load control for shp62, which does not match the figure and was incorrectly labeled as β -actin (top panel below).





- 2.6e. *Relevant interview questions/comments.* There was some confusion as to who performed the experiments for this figure. Dr. Shi thought that Dr. Kim did it (Appendix 052 (Shi interview, page 35)), whereas Dr. Wang said that he did (Appendix 058 (Wang interview, pages 25-26)). Dr. Zhang indicated that Dr. Wang did all experiments for figure 3 (Appendix 054 (Zhang interview, page 32)). Dr. Shi indicated that he does not know what was used for loading control in the figure and thinks that actinin is an error (Appendix 052 (Shi interview, pages 52-53)). Dr. Zhang indicated that Dr. Kim mislabeled actinin as beta-actin (Appendix 054 (Zhang interview, page 58)). In Exhibits K and L to Dr. Wang's interview (Appendix 059), he showed that he performed both actin and actinin loading controls. He further stated that he forgot to properly label the actinin in the final published figure (Appendix 058 (Wang interview, pages 52-56)).
- 2.6f. *Conclusion for this specific example.* Based on the data provided and the interviews, appropriate loading control experiments were performed, although they were labeled incorrectly in the published figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification and fabrication.
- 2.7. M8: Wang, L., Kung, L., Hiltron, J.A., Son, Y.O., Wang, X., Budhraj, A., Lee, J.C., Pratheeshkumar, P., Chen, G., Zhang, Z., Luo, J., and Shi, X. Apigenin suppresses migration and invasion of transformed cells through down-regulation of C-X-C chemokine receptor 4 expression. *Toxicology and Applied Pharmacology* 272, 108-116, 2013 (Appendix 020): Fig. 4C.
- 2.7a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 2.7b. *Original figure from Appendix 020, page 113.*

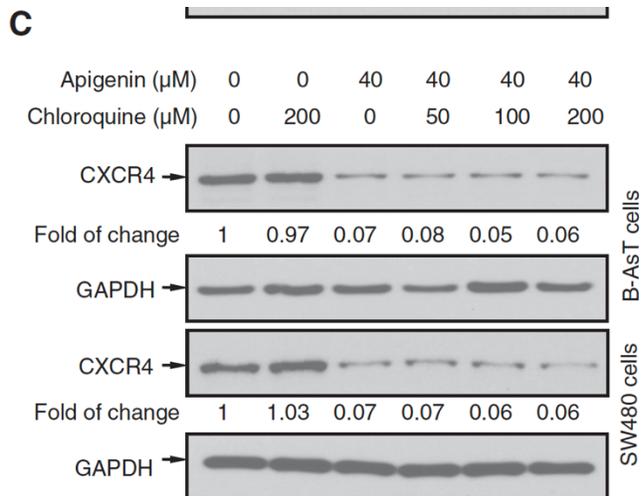
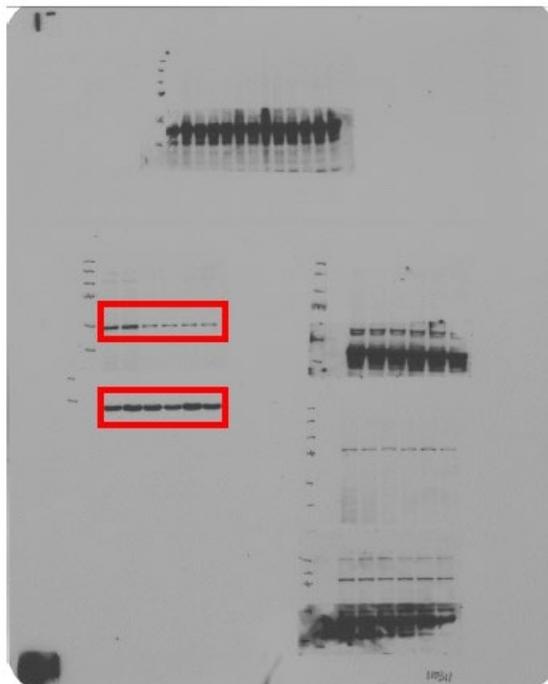


Fig. 4. Down-regulation of CXCR4 by apigenin is not mediated through protein degradation. (A) Cells were treated with the indicated concentration of MG132 for 1 h at 37 °C, followed by treatment with 40 μM apigenin for 24 h. Western blot analysis with antibodies against CXCR4 and GAPDH was performed. (B) Cell lysates were incubating with CXCR4 antibodies overnight at 4 °C. Immunocomplexes were collected through Protein G Agarose. Western blot assay was applied probed with anti-ubiquitin antibody. (C) Cells were treated with the indicated concentration of chloroquine for 1 h at 37 °C, followed by treatment with 40 μM apigenin for 24 h, and then western blotting was performed as described above.

2.7c. *Data provided.* See Appendix 104.
From the Respondents:



Document name: 110311 004

2.7d. *Our analysis.* The data provided for the load controls (GAPDH-second strip from the top; 110311004 in Shi11-02-18/Wang L-TAAP 2016-Fig4c/Wang L TAAP-2013.PDF) are not what was included in the published figure.

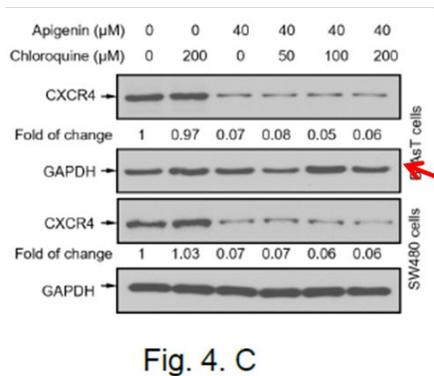
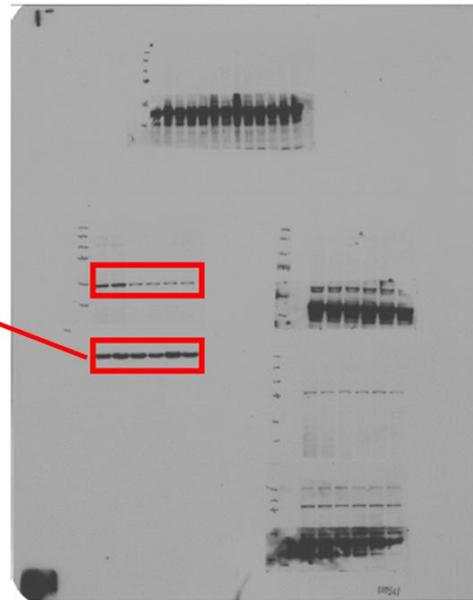


Fig. 4. C



Document name: 110311 004

2.7e. *Relevant interview questions/comments.* Not applicable.

2.7f. *Conclusion for this specific example.* The data provided for the load controls are not what was included in the published figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

2.8. M7: Yin, Y., Li, W., Son, Y.O., Sun, L., Kim, D., Wang, X., Yao, H., Wang, L., Pratheeshkumar, P., Hitron, A., Luo, J., Gao, N., and Shi, X., and Zhang, Z. Quercitrin protects skin from UVB-induced oxidative damage. *Toxicology and Applied Pharmacology* 269, 89-99, 2013 (Appendix 19: Figs. 1E and 5A)

2.8a. *Date requested:* October 24, 2018 by the committee *via* email.

Date received: November 2, 2018 by Respondents *via* flash drive.

2.8b. *Original figure from Appendix 019, pages 92 and 95.*

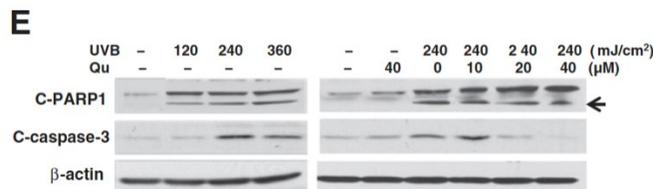


Fig. 1. Effect of quercitrin on apoptosis and apoptotic proteins induced by UVB exposure. A, B, and E, JB6 cells were pretreated with quercitrin (10, 20, and 40 μM) or acetone for 1 h prior UVB exposure. After 24 h, the cells were collected for apoptosis analysis using flow cytometry (A and B) or for immunoblotting assay (E). C and D, 6–8 week old female SKH-1 mice dorsal skin was topically administrated with either acetone (control group) or quercitrin one day before UV exposure. The mice were then exposed to 75 mJ/cm² of UVB for 3 times per week up to 6 weeks. At 1 day and 7 days, the animals were euthanized and dorsal skin tissues were isolated and subjected for immunofluorescence staining of apoptosis (C and D). The positive cells were counted from a total of 500 cells from 8 random field using Olympus BX51 microscope. The results were expressed as a percentage of TUNEL-positive cells (apoptosis index). F, The same as C and D, but mouse dorsal skin tissues were isolated and total protein was extracted for examination of expression levels of cleaved PARP-1 and cleaved caspase-3. * indicates a significant difference compared with control without UVB exposure ($p < 0.05$). # indicates a significant difference compared with 240 mJ/cm² UVB exposure ($p < 0.05$).

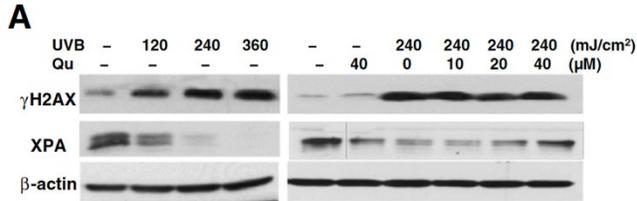
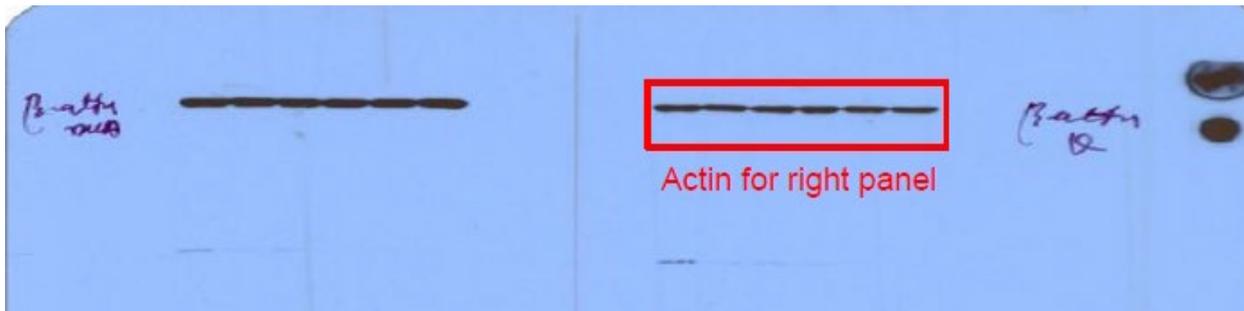


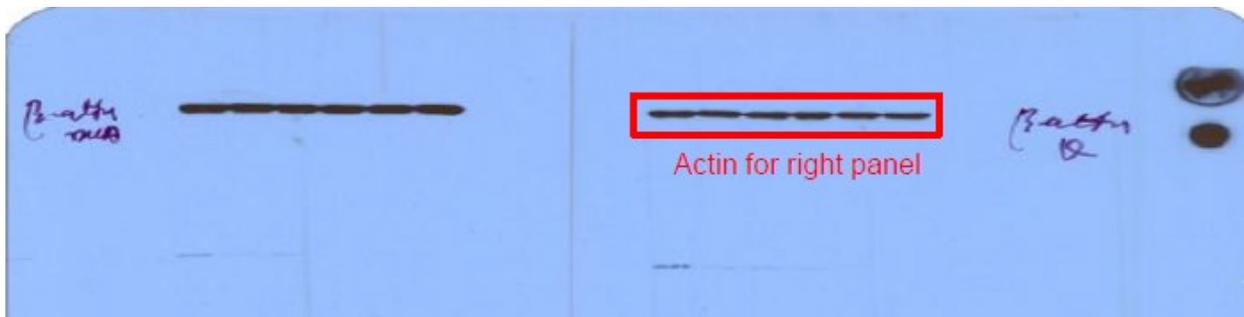
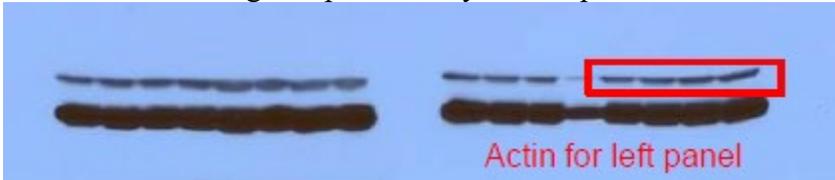
Fig. 5. Effect of quercitrin on DNA repair genes induced by UVB exposure. A, JB6 cells were pretreated with quercitrin (10, 20, and 40 μ M) or acetone for 1 h prior UVB exposure. After 2 h, the cells were collected for immunostaining analysis. B and C, The animal treatment is the same as that in Fig. 1. At 1 day and 7 days of 6-week treatment, the mice were euthanized and mouse skin tissues were isolated. The protein was lysated for immunoblotting analysis (B). The mouse skin sections were used for immunohistological staining with γ H2AX (C).

2.8c. *Data provided.* See Appendix 108.

Load control for Fig. 1E provided by the Respondents:



Load control for Fig. 5A provided by the Respondents:



2.8d. *Our analysis.* The same β -actin loading control was used in Figs. 1E and 5A. However, according to the figure legend, the samples used in Fig. 1E were harvested 24 h post UVB exposure, while those used in Fig. 5A were collected 2 h post UVB exposure. An image of

the same piece of film was provided by the Respondents for the β -actin loading control used in Figs. 1E and 5A. A JPG was provided that was called “Fig 1E-C-actin-LeftJPG”, but no JPG files were provided for Fig. 5A. Unfortunately, as there are no labels/annotations/dates on the blots, it is impossible to verify with which experiments the control really belongs.

- 2.8e. *Relevant interview questions/comments.* Not applicable.
- 2.8f. *Conclusion for this specific example.* Although the experiment in Figs. 1E and 5A were described as different experiments in the legends, the actin loading control was identical. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

3. Inappropriate modification of original data:

Description: The committee defined this category as one-dimensional stretching of original gel images, figure/lane grafting without indication, and/or cropping to remove potentially relevant data.

Overall summary: 3 figures from grants and 3 figure panels from manuscripts. (sections 3.1-3.6).

Overall summary of interviews and relevant comments: During the interviews, we asked each Respondent, as well as two other laboratory personnel (Drs. Hitron and Wang), a series of questions about what image manipulations are considered appropriate when constructing a figure (*i.e.*, contrast/brightness, altering image proportions, splicing lanes from different gel regions, image cropping, color balance, FACS gating). Dr. Shi stated that he does not tell students specifically what the acceptable parameters for changing images are, he just uses “professional common sense” when viewing the images to evaluate them (Appendix 052 (Shi interview, page 23)). Dr. Shi also stated “you can do anything you want as long as in the second one you make it consistent” (Appendix 052 (Shi interview, page 74)). Dr. Shi further stated that data can be manipulated in one or more dimensions as long as when you put it together you make everything consistent (Appendix 052 (Shi interview, page 75)). Additionally, Dr. Shi stated that you can cut blots for figures “whatever you want as long as it is run at the same time”. He stated that it was acceptable to cut lanes out of blots as long as you did not cut blots in from different studies (Appendix 052 (Shi interview, pages 77-78)). Dr. Zhang however only allows manipulation of brightness and contrast and does not allow resizing (Appendix 054 (Zhang interview, pages 22, 67)). Dr. Zhang does not think that data manipulation is appropriate and states “it is not appropriate in her lab to pull data up on one axis” (Appendix 054 (Zhang interview, page 66)). Dr. Hitron stated there were never any standards about data manipulations in Shi lab, and he would only vary brightness and contrast (Appendix 051 (Hitron interview, page 30)). Dr. Hitron further stated “everyone would do their figures a different way, some would scan the whole blot, some would cut out bands, there was no training on how to do figures” (Appendix 051 (Hitron interview, page 31)). Sometimes Dr. Shi would give directions in a lab meeting on the right format to present a figure (Appendix 051 (Hitron interview, page 32)). Dr. Wang stated when he cuts figures from

different data he resizes the parts to make them all the same size (Appendix 058 (Wang interview, pages 68-69)).

Overall conclusions: There were significant inconsistencies in statements from Respondents regarding what data manipulations were allowed in constructing figures for grants and publications. There were no laboratory standards for image manipulation. The committee found multiple incidents of unidimensional image stretching/compressing, inappropriate image cropping, and undocumented image grafting.

Specific examples for “inappropriate modification of original data”:

- 3.1. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 6
- 3.1a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 3.1b. *Original figure from Appendix 008 page 53.*

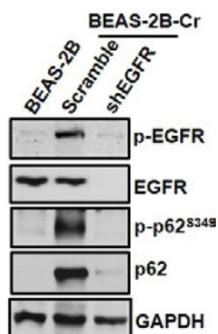
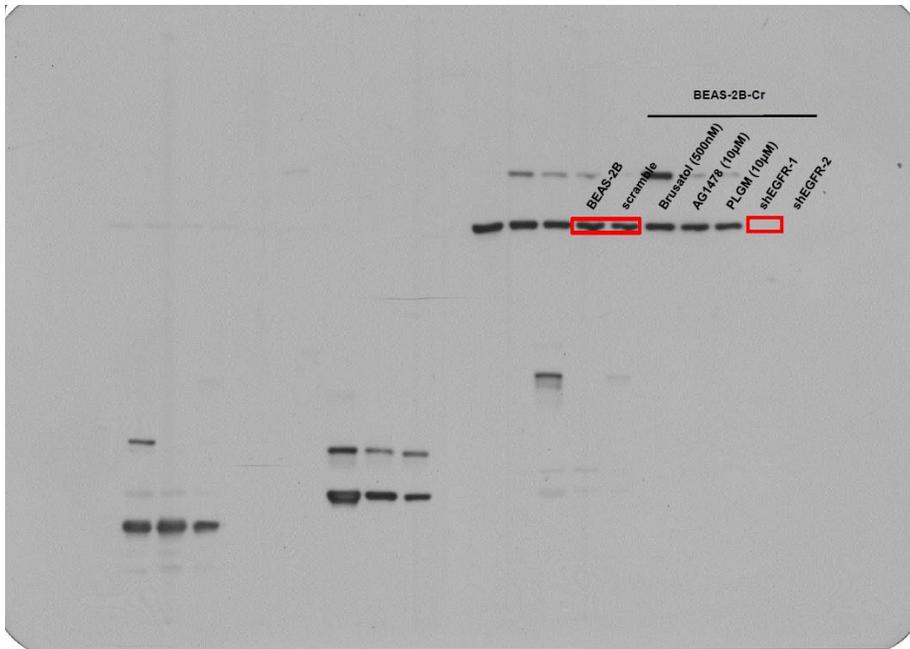
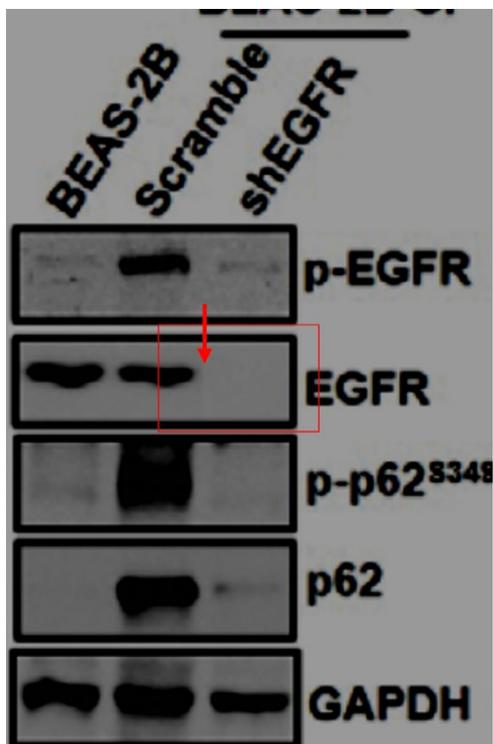


Fig. 6 EGFR upregulates p62 in Cr(VI)-transformed cells. Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) with (shEGFR) or without (Scramble) EGFR knockdown were cultured in 10-cm culture dishes. The cells were harvested and whole protein was isolated for immunoblotting analysis. The results are representative of three independent experiments.

- 3.1c. *Data provided.* See Appendix 109.
From the Respondents:



- 3.1d. *Our analysis.* By visual inspection of the EGFR blot, the committee noticed an unexplained white gap between lanes 2 and 3 (pointed to by a red arrow and highlighted by red box in the figure below). Based on the original data provided by the Respondents, lanes 1 and 2 were grafted onto a **blank** lane of the original gel provided to construct the figure. The lanes used for the figure were marked by the Respondents with red boxes (the full western blot film provided by the Respondents in the figure above). Although it would have been expected, no indication of this grafting was recorded in the figure (*i.e.*, a black line). Moreover, the 3rd lane being blank was an important control for the experiment. The committee could not determine whether this 3rd lane represented authentic data (*i.e.*, whether a sample was loaded in the lane).



- 3.1e. *Relevant interview questions/comments.* Dr. Shi stated that it is acceptable to cut blots for figures “whatever you want, as long as it is run at the same time”. He stated that it was acceptable to cut lanes out of blots as long as you did not cut blots in from different studies (Appendix 052 (Shi interview, pages 77-78)). Dr. Zhang stated that Kim provided the PowerPoint for this figure and labeled it (Appendix 054 (Zhang interview, page 70)).
- 3.1f. *Conclusion for this specific example.* This is an example of taking lanes from different regions of a gel and grafting them together to make it look like they are contiguous. Standard professional practice requires that such grafting be clearly indicated with a solid line or individual box (reference the JBC guidelines in Appendix 110). This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is supported by a preponderance of evidence. The committee concluded that this is a case of data falsification.
- 3.2. G6: Grant Zhang_3200001638 = 1R01ES028984-01 (12/15/2017-11/30/2022) (Appendix 011): Fig. 3A
- 3.2a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 3.2b. *Original figure from Appendix 011, Research Strategy section.*

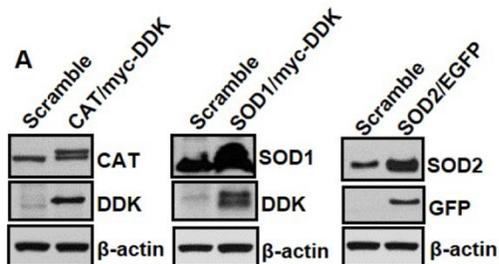


Fig. 3 ROS play a major role in Cd(II)-induced cell transformation. (A) CAT-Myc-DDK-, SOD1-Myc-DDK-, or SOD2-EGFP-tagged plasmid was transfected into BEAS-2B cells followed by antibiotic selection for 1 month to establish stably expressing cells. The cells were harvested and whole protein lysates were isolated. Levels of catalase (CAT), SOD1, and SOD2 were examined using immunoblotting analysis. The results are representative of three independent experiments. (B) BEAS-2B cells with or without stably expressing antioxidant were exposed to 2 μ M of Cd(II) for 2 months. Cell transformation (Soft agar) assay was performed as described in Fig. 2. *, $p < 0.05$ compared to scramble cells without Cd(II) exposure. # and ##, $p < 0.05$ and $p < 0.01$ compared to scramble cells exposed to Cd(II).

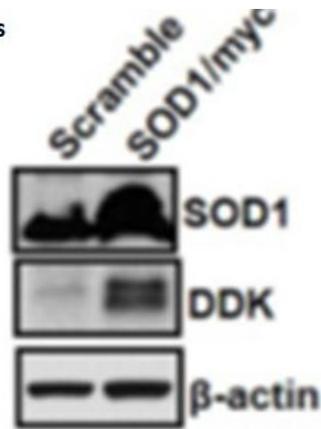
3.2c. *Data provided.* See Appendix 111.

From the Respondents:



3.2d. *Our analysis.* Based on the data provided by the Respondents, the SOD1 blot was significantly compressed on the y-axis to make the original data look like the data in the figure.

Data Provided by Respondents



Y-Axis Compression of the Original Data provided by the Respondents, applied by the committee, to recreate the figure in the grant

Figure in Grant

- 3.2e. *Relevant interview questions/comments.* The committee asked Drs. Shi and Zhang about this figure. Aside from some general comments about figure production (noted in summary above), Dr. Shi had no specific comments. Dr. Zhang indicated that she got the data from Dr. Son, who is currently in Korea. Dr. Zhang does not have the raw data for this figure, and she thinks Dr. Son left this raw data with Dr. Shi, but she indicated she would have to check. Dr. Son did not take the film with him, so he probably provided a scanned image to Dr. Shi (Appendix 054 (Zhang interview, pages 64-65)).
- 3.2f. *Conclusion for this specific example.* This is an example of extreme unidimensional (along the y-axis) image compression. This results in a significant change in the appearance of the data in the grant figure. Specifically, it made tight “bands” appear when there were only diffuse “blobs” in the original data. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.
- 3.3. G7: Grant Zhang_3200001897 = 1R01CA228236-01A1 (01/01/2018-05/31-2023) (Appendix 012): Fig. 12, which is the same as Fig. 3G in M1 Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).
- 3.3a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 3.3b. *Original figure from Appendices 012 page 65, and 013, page 395.*

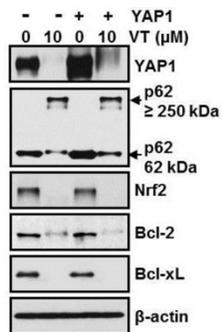


Fig 12 Verteporfin inhibits p62 and its downstream proteins independent of YAP1. PC3 cells were transiently transfected with pCMV6-GFP/YAP1 followed by treatment with or without verteporfin. At 24 h, whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments.

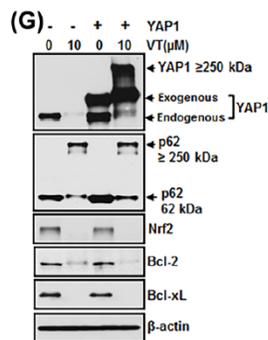
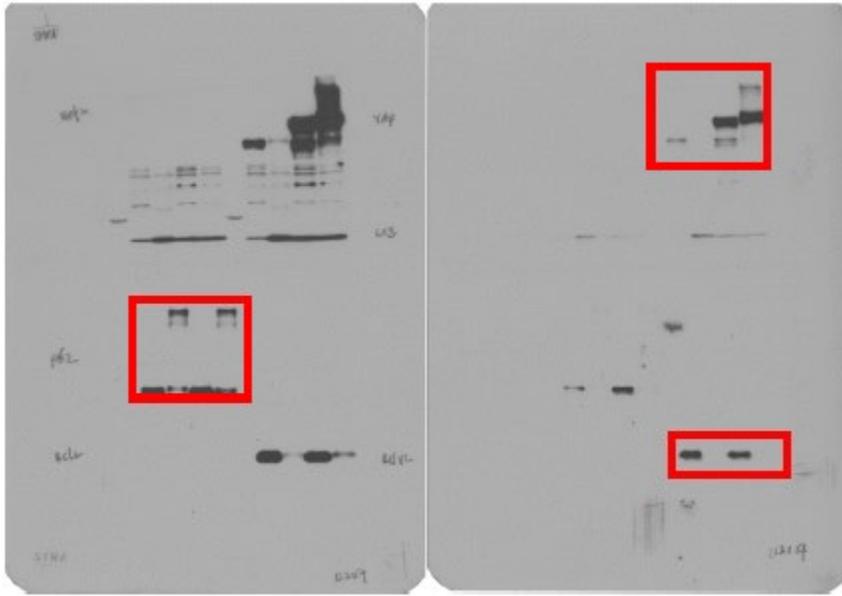


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μ M Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μ M and 10 μ M Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μ M 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean \pm SE (n = 6). * and #, P < 0.05 compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μ M of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

3.3c. *Data provided.* See Appendix 112.

From the Respondents:



File name: Yap, p62, Bcl-xl

- 3.3d. *Our analysis.* The committee noticed that Fig. 12 in the grant G7 and Fig. 3G in manuscript M1 show data from the same experiment (Appendices 012 and 013). When the committee examined the data provided for Fig. 3G and compared it to the published figure and to Fig. 12 in the grant G7, it noticed that the same images were used to create both figures. However, in the grant figure (Fig. 12), the YAP1 portion of the image was seriously cropped to omit relevant higher molecular weight bands and was elongated in the y-axis (figure below). It should be noted that this data manipulation was only obvious upon examination of the original data and would not have been detected by examining the published and grant figures.

Fig. 12 in G7

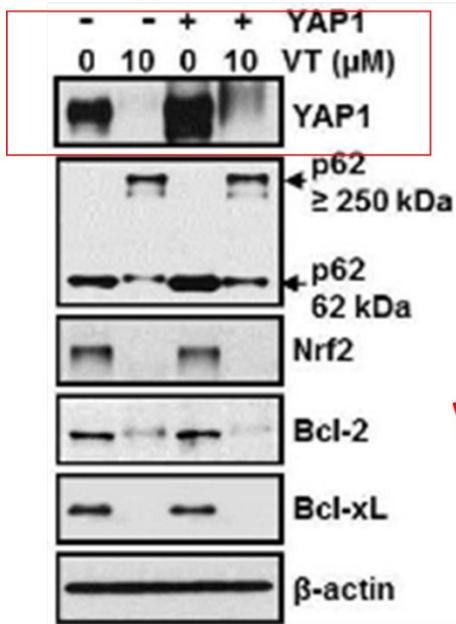
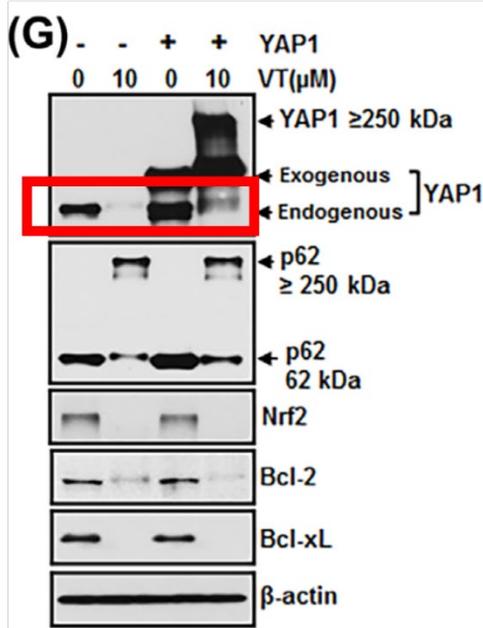


Fig. 3G in M1



Vs.

3.3e. *Relevant interview questions/comments.* Dr. Shi indicated that Dr. Kim “did figure 3G” (Appendix 052 (Shi interview, page 35)).

3.3f. *Conclusion for this specific example.* This is an example of inappropriate cropping and unidimensional stretching of a blot. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

3.4. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 3A

3.4a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

3.4b. *Original figure from Appendix 013, page 395.*

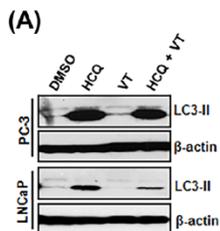
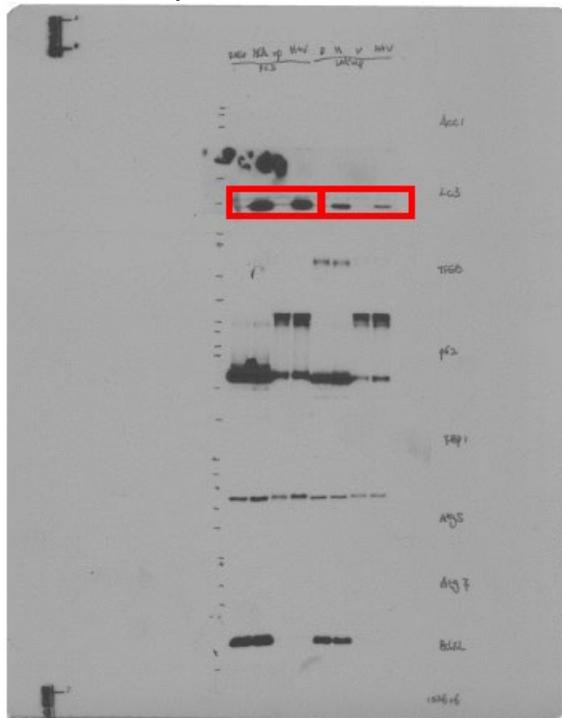


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μ M Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μ M and 10 μ M Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μ M 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean \pm SE ($n = 6$). * and #, $P < 0.05$ compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μ M of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

3.4c. *Data provided.* See Appendix 113.
From the Respondents:



File name: LC3

3.4d. *Our analysis.* Based on the original data provided, the western blot images for LC3-II in the upper and lower half of the figure were stretched along the x-axis to make the final figure (comparison in the figure below).

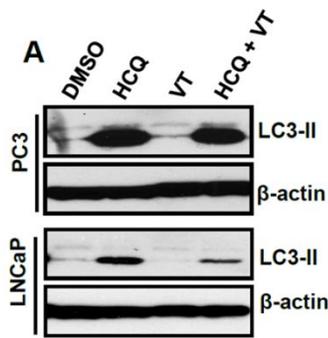


Fig 3

Figure 3A raw data provided by the Respondents

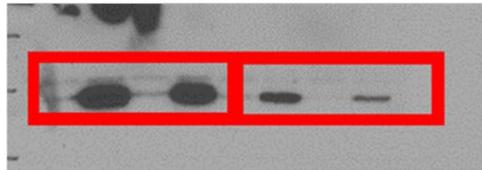
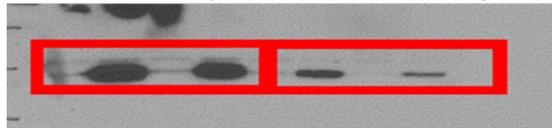


Figure 3A raw data provided by the Respondents and manipulated by the committee to elongate



3.4e. *Relevant interview questions/comments.* Not applicable.

3.4f. *Conclusion for this specific example.* This is an example of unidimensional (along the x-axis) image stretching. This results in a change in the appearance of the data in the grant figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

3.5. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 5E

3.5a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

3.5b. *Original figure from Appendix 013, page 397.*

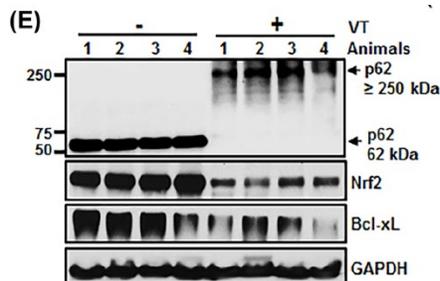
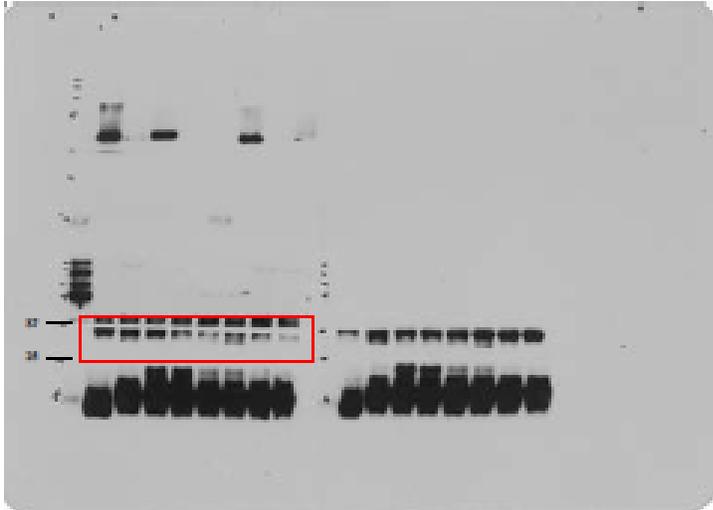


FIGURE 5 p62 increase in tumor growth and Verteporfin inhibition. (A-E) 6-week-old male immunodeficient nude mice were s.c. injected with 1.5×10^6 PZ-HPV7 cells with or without p62 expression into right or left flank, respectively. At 2 weeks post-injection, the animals were *i.p.* administrated with either 100 mg/kg Verteporfin (VT) or 10% DMSO (control), once per day, three times per week. At five weeks post-injection, the animals were euthanized. Pictures were captured in the animals injected with (right) or without (left) p62-expressing cells (A). Tumor volumes ($(\text{Length} \times \text{Width}^2)/2$) were measured (B). Tumors were isolated from the animals injected with p62-expressing PZ-HPV7 cells with or without Verteporfin treatment (C-E). The images of those tumors were captured (C) and tumor volume was measured (D). (E) Whole-protein lysates were collected from tumor tissues and used for immunoblotting analysis. The results are shown as mean \pm SE ($n = 4$). $*P < 0.05$, compared to scramble cells (B) or p62-expressing cells with and without treatment (D). (F-H) 6-week-old male immunodeficient nude mice were randomly divided into two groups with 8 animals/group. 5×10^6 PC-3 cells were s.c. injected into both flanks of each animal. The animals were administrated with either 100 mg/kg Verteporfin (VT) or 10% DMSO (control), once per day, three times per week. At 2 weeks post-injection, animals were euthanized and tumors were isolated. The images of those tumors were captured (F) and tumor weight was measured (G). (H) Wholeprotein lysates were collected from tumor tissues and used for immunoblotting analysis. The results represent three-independent experiments

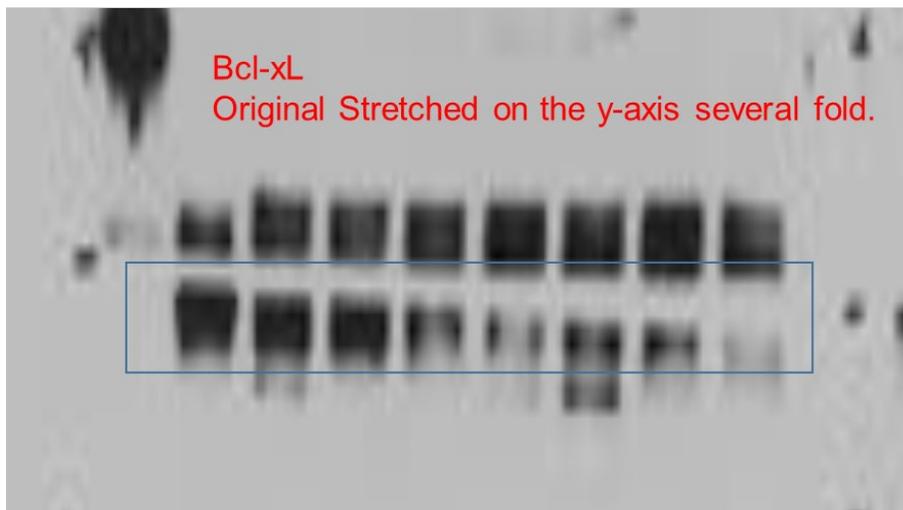
3.5c. *Data provided.* See Appendix 114.

Bcl-xL



[dki/total/experiment/data/WB/2017/073017/3/scan3](#)

3.5d. *Our analysis.* Based on the data provided, the original western blot image for Bcl-xL was significantly stretched on the y-axis to make the published figure (figure below). This manipulation changed the appearance of the data since the upper band of the doublet in the original figure was not represented in the published data.



- 3.5e. *Relevant interview questions/comments.* Not applicable.
- 3.5f. *Conclusion for this specific example.* This is an example of extreme unidimensional (along the y-axis) image compression. This results in a significant change in the appearance of the data in the grant figure. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.
- 3.6. M7: Yin, Y., Li, W., Son, Y.O., Sun, L., Kim, D., Wang, X., Yao, H., Wang, L., Pratheeshkumar, P., Hitron, A., Luo, J., Gao, N., and Shi, X., and Zhang, Z. Quercitrin protects skin from UVB-induced oxidative damage. *Toxicology and Applied Pharmacology* 269, 89-99, 2013 (Appendix 019) Fig. 5A
- 3.6a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 3.6b. *Original figure from Appendix 019, page 95.*

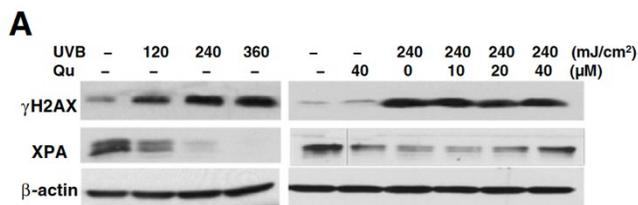
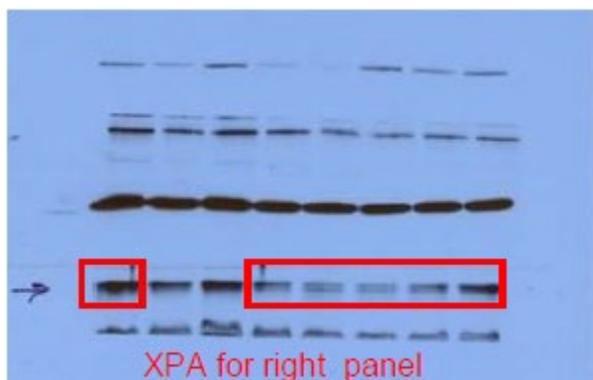


Fig. 5. Effect of quercitrin on DNA repair genes induced by UVB exposure. A, JB6 cells were pretreated with quercitrin (10, 20, and 40 μ M) or acetone for 1 h prior UVB exposure. After 2 h, the cells were collected for immunostaining analysis. B and C. The animal treatment is the same as that in Fig. 1. At 1 day and 7 days of 6-week treatment, the mice were euthanized and mouse skin tissues were isolated. The protein was lysated for immunoblotting analysis (B). The mouse skin sections were used for immunohistological staining with γ H2AX (C).

- 3.6c. *Data provided.* See Appendix 115.



- 3.6d. *Our analysis.* By visual inspection of the right panel of the XPA blot, the committee noticed an unexplained white gap between lanes 1 and 2. Based on the original data provided by the Respondents, lane 1 was grafted on to the fourth lane of the original gel provided to construct the published figure (figure above). The lanes used for the figure were marked

by the Respondents with red boxes. Although it would have been expected, no indication of this grafting was recorded in the figure (*i.e.*, a black line).

- 3.6e. *Relevant interview questions/comments.* Not applicable.
- 3.6f. *Conclusion for this specific example.* This is an example of taking lanes from different regions of a gel and grafting them together to make it look like they are contiguous. Standard professional practice requires that such grafting be clearly indicated with a solid line or individual box (reference the JBC guidelines in Appendix 110). This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification and fabrication.

4. Inappropriate scale bars on images:

Description: The committee defined this category as scale bars with the wrong units, that could not be confirmed with metadata, and/or seemed obviously wrong based on what was imaged.

Overall summary: 4 figures from grants and 1 figure panel from manuscripts. (sections 4.1-4.6).

Overall summary of interviews and relevant comments: The committee observed (on 10/19/2018) Dr. Kim taking images with the laboratory microscope and then using those images to construct a figure with a scale bar. The scale bar was added manually using calculations that the committee did not completely understand. Dr. Kim did not rely on the imaging software. (Appendix 035)

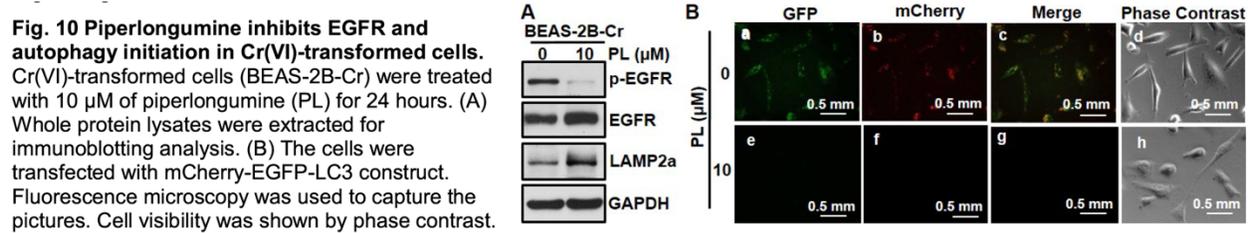
From the interview: Dr. Shi stated that Dr. Kim assigns the scale bars to the images. In response to several questions about scale bars, Dr. Shi stated that he cannot say whether the scale bars are wrong because he is “not good at that kind of stuff” (Appendix 052 (Shi interview, pages 42-46)). Dr. Zhang stated that she thinks the scale bars are inserted automatically, and she relies on Dr. Kim for this (Appendix 054 (Zhang interview, pages 44-45)).

Overall conclusions: The committee saw several incidences where the scale bars used in microscopy images were suspect. In some cases, they were the wrong unit (*i.e.*, μM instead of the correct μm). In other cases, the magnification of the original images was not consistent with the scale bar used in the final figures or with the cell or tissue being imaged. For reference the committee is using a size range of 2-10 μm for the size of an average mammalian cell nucleus as its basis for evaluating images. The use of PowerPoint templates for figure construction from the Shi and the Zhang laboratories, without verifying the images or units, promotes the types of mistakes that the committee observed. Based on the committee’s interactions with the Respondents and the numerous errors in their figures, the Respondents did not appear to care about the accuracy of the scale bars used in their figures.

Specific examples for “inappropriate scale bars on images”:

- 4.1 G1: Grant Shi_3210000529 = 3R01ES025515-03S1 (02/01/2017-01/30/2020) (Appendix 006, pages 11-133): Fig. 10B
- 4.1a. *This figure was originally provided to the committee at the start of the investigation.*

4.1b. *Original figure from Appendix 006 page 63 of the research strategy section of the grant.*



4.1c. *Data provided.* See inquiry report (Appendix 006, pages 11-133). In addition, Dr. Shi, after the interviews, provided some file paths for Fig. 10 (Appendix 065).

4.1d. *Our analysis.* Based on the scale bars in the figure (0.5 mm which equals 500 μ m), the transformed cells are up to 1mm and the nuclei are up to \sim 200 μ m. These are not valid sizes for the types of fibroblastoid cells shown in the images.

4.1e. *Relevant interview questions/comments.* See general comments in the summary above for Section 4.

4.1f. *Conclusion for this specific example.* Based on the cells being imaged, the scale bars recorded in the figure are incorrect. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

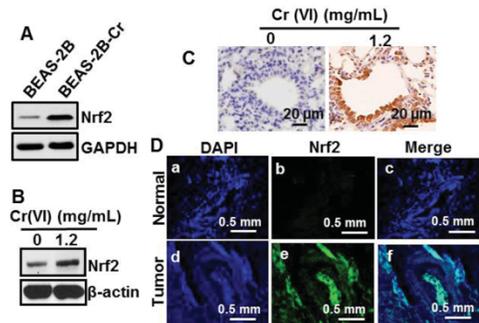
4.2.1 G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 7D

4.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

4.2b. *Original figure from Appendix 008 page 54.*

Fig. 7 Increased Nrf2 protein level in Cr(VI)-transformed cells, in lung tissues from Cr(VI)-exposed animals, and in lung tumor tissues from a worker exposed to Cr(VI) for 19 years. (A) Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) were cultured in 10-cm dishes. Whole protein was isolated for examination of Nrf2 level using immunoblotting. The results are representative of three independent experiments. (B) and (C) 4-week old both male and female BALB/c mice were intranasally exposed to Cr(VI) particles (1.2 mg/mL) as described in Fig. 2. Lung tissues were isolated to examine Nrf2 level using immunoblotting (B) and immunohistological (C) analyses. Intensity of brown color represents Nrf2 level (C). Sections are representative of 6 different animals from each treatment group. (D) Formalin-fixed human lung tumor and adjacent normal tissues from a non-smoking worker exposed to Cr(VI) were subjected to fluorescence immunostainings with DAPI (nuclear control, a and d) and Nrf2 (b and e). The merged images are provided in c and f. The results are representative of three independent experiments.



4.2c. *Data provided.* See Appendix 116.

4.2d. *Our analysis.* Based on the images provided, only a tiny fraction of the original images was used in Fig. 7D. The original scale bar in the data for panels 7Da,b,c were 50 (see

example above) μm and all the scale bars in the grant figure were 0.5 mm (500 μm). This 10-fold difference is not consistent with either set of images.

4.2e. *Relevant interview questions/comments.* Not applicable.

4.2f. *Conclusion for this specific example.* The scale bars used in the Figure and those in the provided data do not match. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

4.3 G5: Grant Zhang_3200001472 = 5R01ES028321-02 (09/01/2017-07/31/2022) (Appendix 010): Fig. 11 panel O

4.3a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

4.3b. *Original figure from Appendix 010 page 33.*

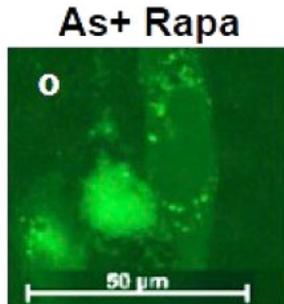
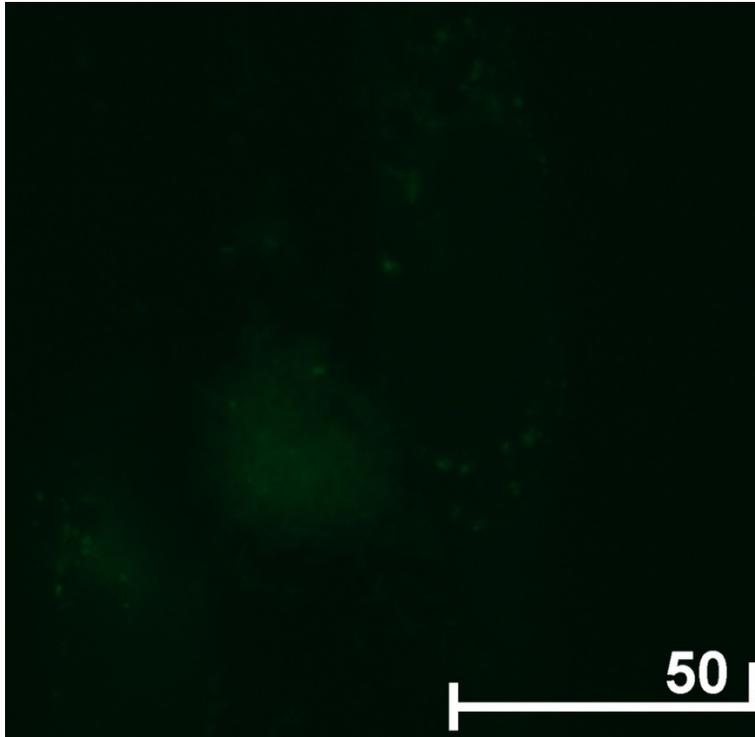


Fig. 11 Autophagy impairment in arsenic-transformed cells. Passage-matched normal cells (BEAS-2B) and arsenic-transformed cells (BEAS-2B-As) were transfected with mCherry-GFP-LC3 plasmid and exposed to arsenic (20 μM) and rapamycin (0.1 μM) for 24 hrs. The cells were visualized using fluorescence microscopy. The results are representative of three independent experiments.

4.3c. *Data provided.* See Appendix 117.



- 4.3d. *Our analysis.* The committee examined all the panels in Fig. 11 and all the original data provided by the Respondents. The original images were obviously adjusted for brightness and contrast to create the final images in the grant figure. Scale bars were only recorded in the original images for A, C, D, E, I, J, K, L, N, O (half of the scale bar was cropped out of the image), P, V, W, and X. Original images I, J, K, L, O, V, W, and X were at a different magnification than the other images. Likely these were cropped from an original image that was not provided. Given that the magnified images were cropped, it was not possible to precisely determine their magnification, however, based on the images' pixel size, it appears that they were magnified ~2-fold. The images were arranged in trios (*i.e.*, A, E, I) and in several of these trios the third image was not the same magnification as the first two (*e.g.*, A, E, I; C, G, K; D, H, L; N, R, V; P, T, X). Based on a visual comparison of the grant Figure and the data provided, the scale bars seemed to be consistent.
- 4.3e. *Relevant interview questions/comments.* Not applicable.
- 4.3f. *Conclusion for this specific example.* While scale bars were not present in all the original images, they seemed to be correctly generated. The use of different magnifications of the merged images (generally the third one of the trio) is not a standard way to create images for publication; however, it does seem like, in this case, the scales bars were correct. This is a minor deviation from accepted practices of the research community but the committee was not able to determine whether research misconduct occurred.
- 4.4 G5: Grant Zhang_3200001472 = 5R01ES028321-02 (09/01/2017-07/31/2022) (Appendix 010): Fig. 15
- 4.4a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents via flash drive.

4.4b. Original figure from Appendix 010 page 34.

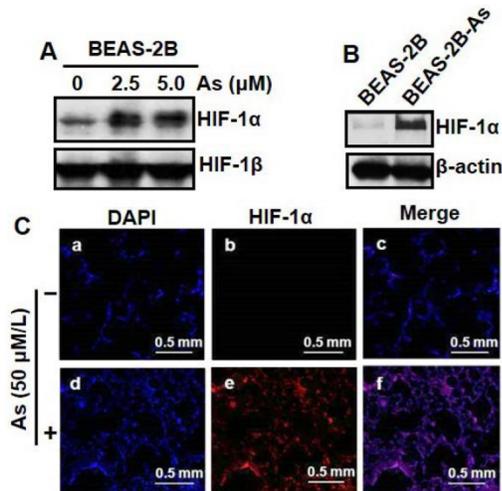
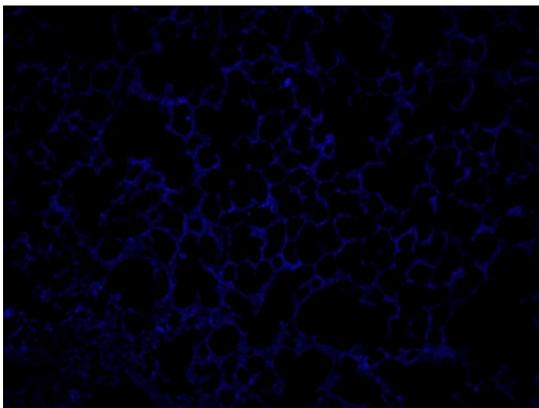


Fig. 15 Arsenic exposure increases HIF-1 α expression in vitro and in vivo. (A) BEAS-2B cells were treated with sodium arsenite for 6 hrs. The whole cell lysates were collected to examine expression of HIF-1 α using immunoblotting. (B) Arsenic-transformed cells (BEAS-2B-As) and passage-matched normal ones (BEAS-2B) were subjected to examination of HIF-1 α expression using immunoblotting. The above immunoblotting results are representative of three independent experiments. (C) Exposure of animals to arsenic is described in Fig. 9. Tissue sections of lung were used to examine expression of HIF-1 α . 4',6-diamidino-2-phenylindole (DAPI) was used for nuclear staining. Sections are representative of six different mice from each treatment group.

4.4c. *Data provided.* See Appendix 118. Comment on the data provided: The TIF file images used in Fig. 15C had a complete set of metadata with exposure times, camera systems and focusing mode. The committee examined this metadata, and the exposure times were similar and consistent with the images.

Fig. 15Ca



4.4d. *Our analysis.* There were no scale bars in the TIF images used to construct the figure panels in Fig. 15C (one example above). From the metadata, it was not possible to generate an accurate scale bar because there was no information on the size (in μm) of the image.

- 4.4e. *Relevant interview questions/comments.* Not applicable.
- 4.4f. *Conclusion for this specific example.* Based on the data provided, it is not possible to determine if the scale bars are accurate. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 4.5 M12: Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROS-dependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Fig. 3E
- 4.5a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 4.5b. *Original figure from Appendix 024, page 5.*

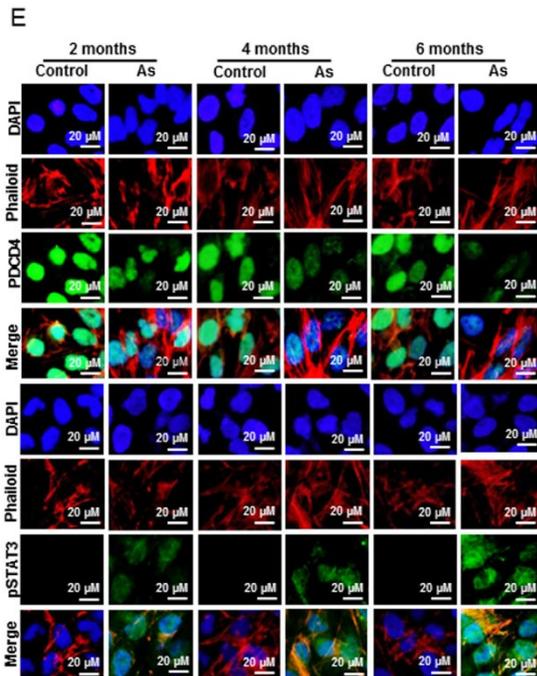
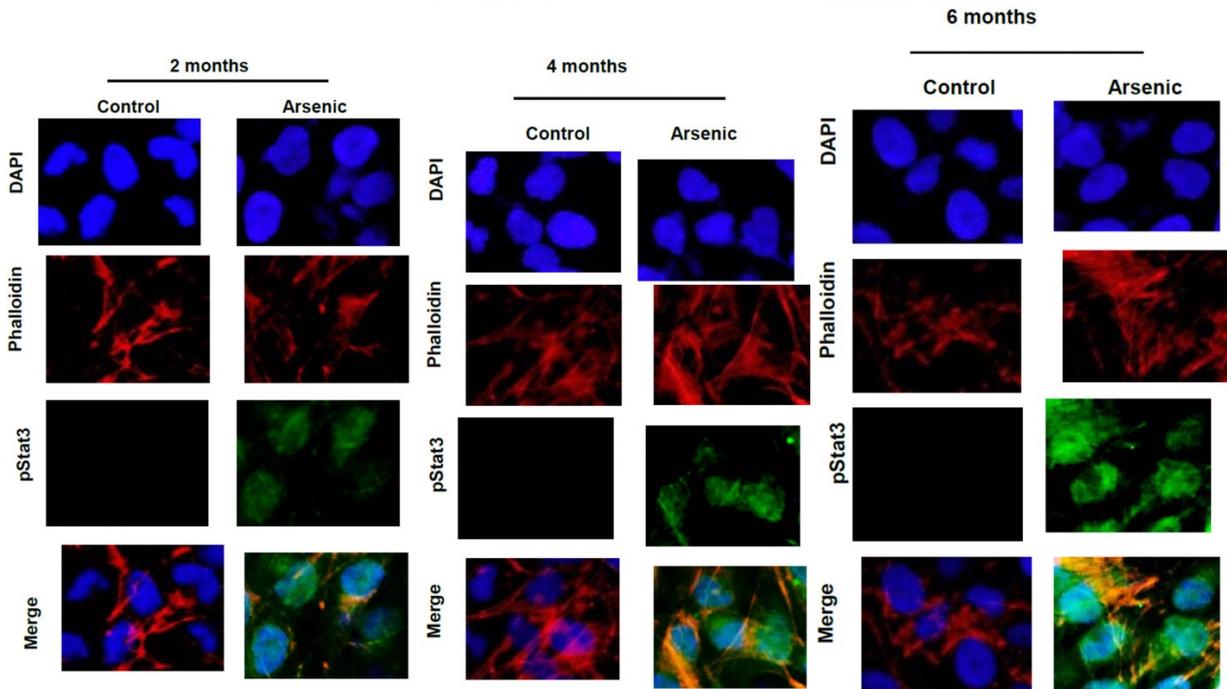
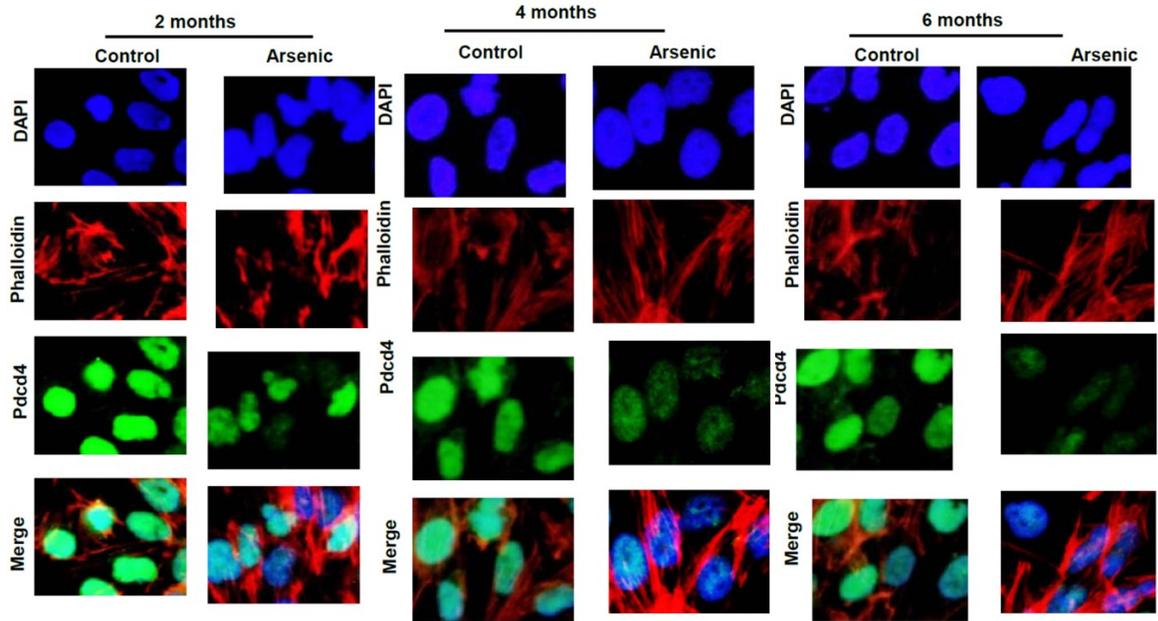


Figure 3. The arsenic-induced miR-21 increase and PDCD4 suppression contribute to malignant cell transformation. BEAS-2B cells were maintained in a medium containing various concentrations of arsenic (0.1, 0.25 and 0.5 μ M) for 6 months. (A,B) Cells were cultured in 0.35% soft agar for 5 weeks and number of colonies in the entire dish counted. (A) Representative images of control (left panel) and arsenic-treated (right panel) colonies. (B) Colony number increased in a dose-dependent manner. (C) The relative miR-21 level, determined by Taqman real-time PCR, increased in a time- and dose-dependent manner. (D) Total cell lysates were prepared for western blot analysis after 2, 4 and 6 months exposure to arsenic using specific antibodies against PDCD4, p47phox, pSTAT3 and STAT3. Apparent protein levels for PDCD4 decreased and P47phox and pSTAT3 increased in a time- and dose-dependent manner. (E) Representative images of fluorescence immunostaining for PDCD4 and pSTAT3 after 2, 4 and 6 months exposure to arsenic, and confirm results from western blot analysis. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *Indicates a statistically significant difference compared to control with $p < 0.05$.

4.5c. *Data provided.* See Appendix 119.



- 4.5d. *Our analysis.* The data provided had no scale bars recorded, so it was not possible to determine the veracity of the scale bars included in the published figure. Additionally, as recorded in the published figure, the scale bars are the incorrect unit, micromolar (μM), when they should be micron (μm).
- 4.5e. *Relevant interview questions/comments.* See general comments in the summary above for Section 4.

4.5f. *Conclusion for this specific example.* This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

5. Data provided did not match the published figure:

Description: The committee defined this category as “The Respondents did not supply the original data that matched the figures in grants or manuscripts”.

Overall summary: 2 figures from grants and 1 figure panel from manuscripts. (sections 5.1-5.3).

Overall summary of interviews and relevant comments: See individual sections below for specific comments.

Overall conclusions: Based on observations of laboratory notebooks and data storage in the laboratory, the Respondents’ inability to provide precise raw data is likely caused by a lack of good data management practices. Additionally, the committee noticed that Dr. Shi’s group uses templates to generate figures, placing new data into old figure templates. The labels on the old figures may not be updated to reflect the new experiment. Compounding this is the fact that Dr. Shi and Dr. Zhang do not regularly check the original data and rely instead on PowerPoint slides produced by staff.

Specific examples for “data provided did not match the published figure”:

- 5.1. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 8B
- 5.1a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 5.1b. *Original figure from Appendix 008 page 54.*

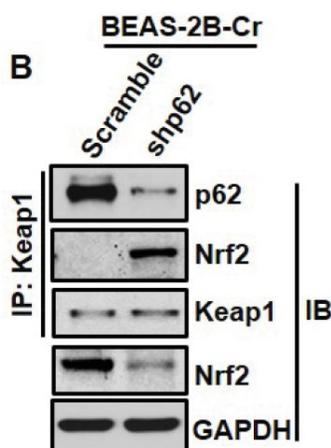
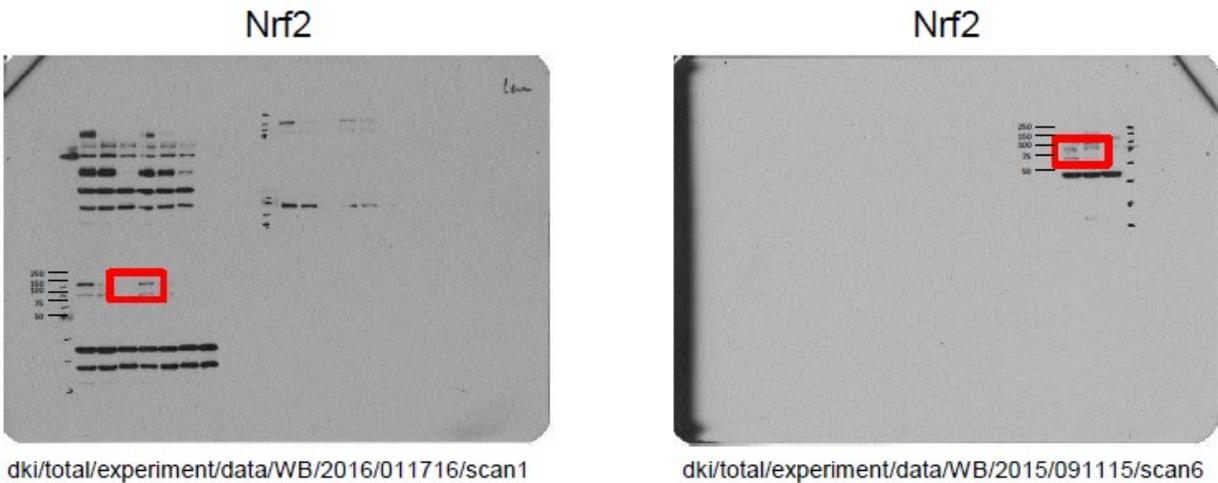
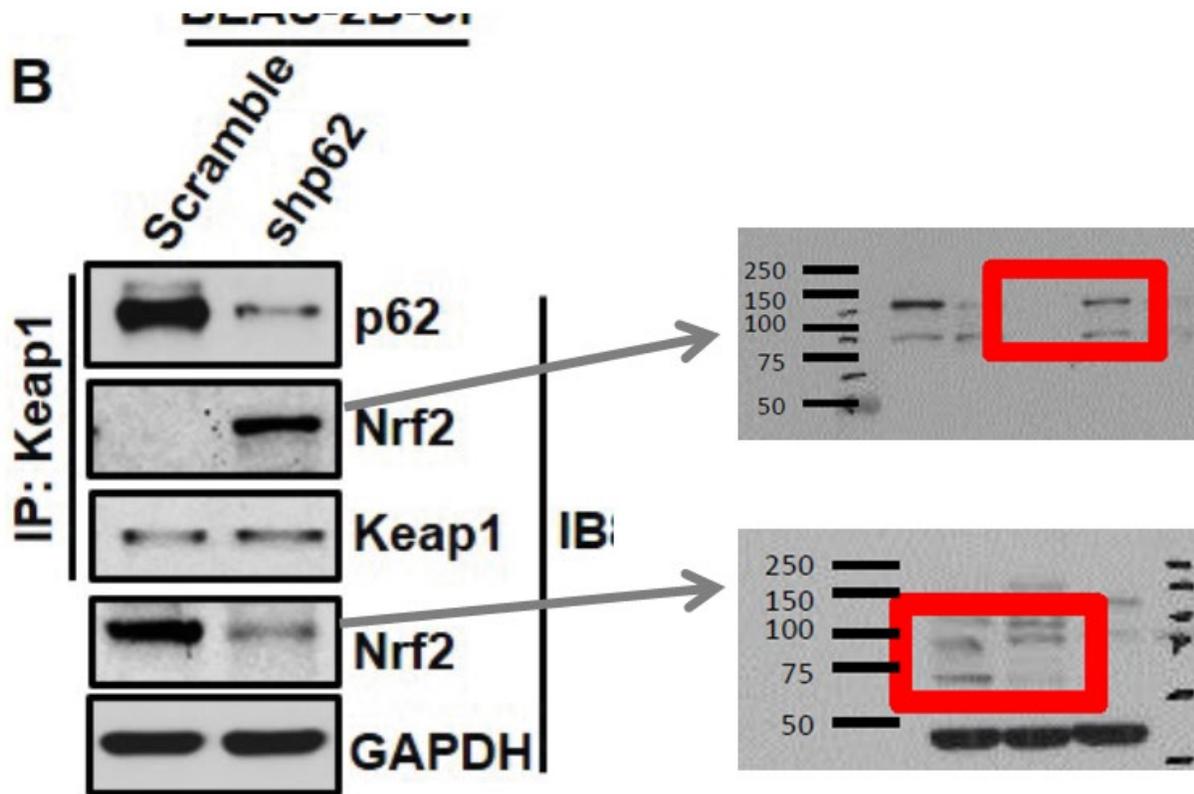


Fig. 8 p62 upregulates Nrf2 in Cr(VI)-transformed cells. (A) and (B) Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) with (shp62) and without (Scramble) knockdown of p62 by its shRNA were subjected to either co-immunoprecipitation and immunoblotting analyses. The results are representative of three independent experiments.

5.1c. *Data provided.* See Appendix 103.



5.1d. *Our analysis.* From the data provided, the blots for Nrf2 did not match the figure in the grant. The brightness and contrast did not match, the background “speckles” in the grant figure did not match the data provided (especially in the upper blot), and the “smudge” on the lower right of the band in the lower Nrf2 panel, “scramble”, did not match the data provided. Additionally, based on the data provided, there was confusion over the molecular weights of the bands shown in the grant figure. In provided data marked “dki/total/experiment/data/WB/2016/011716/scan1”, the bands represent a protein of the incorrect molecular weight (The committee assumed that Nrf2 was a ~68 kDa protein).



5.1e. *Relevant interview questions/comments.* Not applicable.

5.1f. *Conclusion for this specific example.* The data shown for Nrf2, in the grant, cannot be verified from the data provided. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

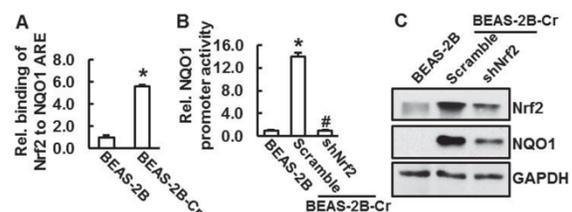
5.2. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 9B

5.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

5.2b. *Original figure from Appendix 008 page 54.*

Fig. 9 Upregulation of NQO1 by Nrf2 in Cr(VI)-transformed cells. (A) ChIP assay and qRT-PCR. Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) were fixed with formaldehyde and cross-linked. The chromatin was sheared and immunoprecipitated with anti-Nrf2 antibody or control IgG. Binding of Nrf2 to NQO1 promoter was analyzed by qPCR using specific primers for ARE of NQO1 promoter. GAPDH was used as a control. (B) and (C) BEAS-2B and BEAS-2B-Cr cells with (shNrf2) or without (Scramble) Nrf2 knockdown were cultured in 10-cm dishes. (B) The cells were transfected with NQO1-ARE firefly luciferase and renilla luciferase. Relative luciferase activity was measured. (A) and (B) Data are mean \pm SD (n=6). * and #, $p < 0.05$, compared to BEAS-2B cells and BEAS-2B-Cr-Scramble, respectively. (C) Whole protein lysates were isolated for immunoblotting analysis. The results are representative of three independent experiments.



- 5.2c. *Data provided.* See Appendix 120.
- 5.2d. *Our analysis.* In the grant Fig. 9B, the third column of the bar graph is labeled “shNrf2”. However, in the data provided by the Respondents (figure in section 5.2c), the third column of this graph is labeled “brusatol”. Along with the data provided, the Respondents inserted a note stating that “There is a typo. shNrf2s should be brusatol”. A search of the grant document for the word “brusatol” did not yield results. Thus, the drug was not mentioned anywhere in the grant.
- 5.2e. *Relevant interview questions/comments.* The committee asked Drs. Shi and Zhang how Fig. 9 was generated. Dr. Shi stated that when the committee requested the information on this figure, the Respondents found a mistake in panel B. There was a miscommunication in the lab which is why there is an error in the figure. Because of this, the figure legend is wrong. The error occurred because they do not use a chemical inhibitor (brusatol) very often. Dr. Shi stated “It does not change the data because the concept is the same” (Appendix 052 (Shi interview, pages 48-49)). Dr. Shi further stated “the error happened because they did not check it before inserting the figure in the grant” (Appendix 052 (Shi interview, page 50)). Dr. Zhang stated that “this is my mistake”. Dr. Kim told her that he used brusatol instead of the knockdown (shNrf2). Dr. Zhang did not check the lab notebook to see what was used. Dr. Kim gave her the figure in PowerPoint (Appendix 054 (Zhang interview, pages 48-49)). Dr. Zhang did not see the original data (Appendix 054 (Zhang interview, page 50)). Dr. Zhang also stated that this error did not change the conclusion of the experiment (Appendix 054 (Zhang interview, pages 52-53)).
- 5.2f. *Conclusion for this specific example.* The committee noticed that Dr. Shi’s group uses templates to generate figures. New data are placed into old figure templates, and the labels may not be changed to reflect the new experiment. Compounding this is the fact that the two PIs do not regularly check the raw data and rely on PowerPoint slides produced by staff. These laboratory practices likely precipitated the error in this figure and perhaps in other figures as well. The misrepresentation described in the section is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.
- 5.3. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013) Fig. 5H
- 5.3a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 5.3b. *Original figure from Appendix 013, page 397.*

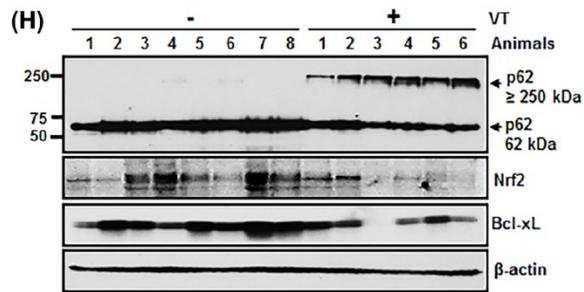
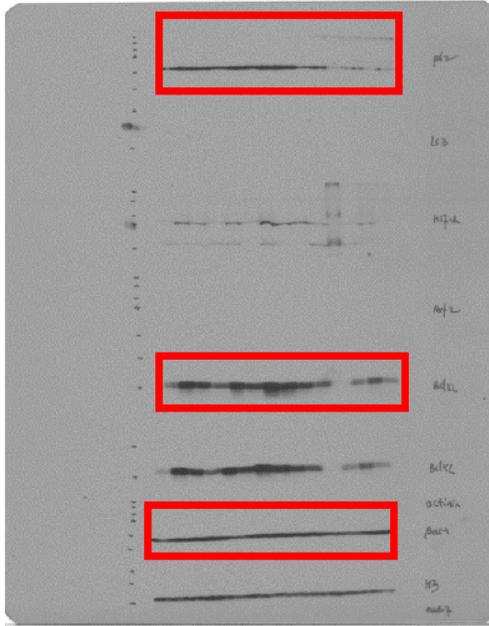
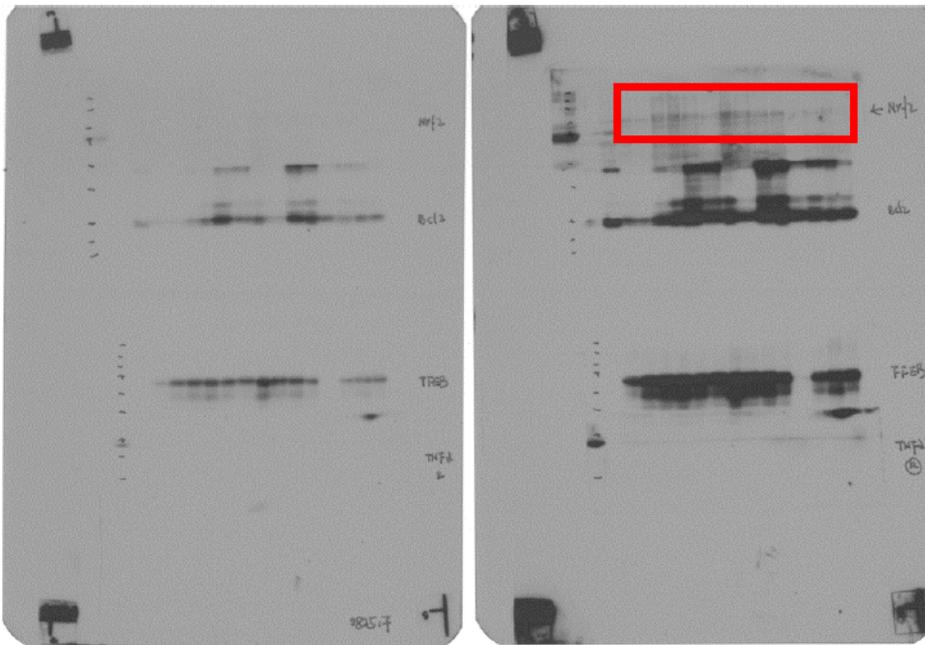


FIGURE 5 p62 increase in tumor growth and Verteporfin inhibition. (A-E) 6-week-old male immunodeficient nude mice were s.c. injected with 1.5×10^6 PZ-HPV7 cells with or without p62 expression into right or left flank, respectively. At 2 weeks post-injection, the animals were *i.p.* administrated with either 100 mg/kg Verteporfin (VT) or 10% DMSO (control), once per day, three times per week. At five weeks post-injection, the animals were euthanized. Pictures were captured in the animals injected with (right) or without (left) p62-expressing cells (A). Tumor volumes ($(\text{Length} \times \text{Width}^2)/2$) were measured (B). Tumors were isolated from the animals injected with p62-expressing PZ-HPV7 cells with or without Verteporfin treatment (C-E). The images of those tumors were captured (C) and tumor volume was measured (D). (E) Whole-protein lysates were collected from tumor tissues and used for immunoblotting analysis. The results are shown as mean \pm SE ($n = 4$). $*P < 0.05$, compared to scramble cells (B) or p62-expressing cells with and without treatment (D). (F-H) 6-week-old male immunodeficient nude mice were randomly divided into two groups with 8 animals/group. 5×10^6 PC-3 cells were s.c. injected into both flanks of each animal. The animals were administrated with either 100 mg/kg Verteporfin (VT) or 10% DMSO (control), once per day, three times per week. At 2 weeks post-injection, animals were euthanized and tumors were isolated. The images of those tumors were captured (F) and tumor weight was measured (G). (H) Wholeprotein lysates were collected from tumor tissues and used for immunoblotting analysis. The results represent three-independent experiments

5.3c. *Data provided.* See Appendix 118.



File name: p62, Bcl-x1, actin



File name: Nrf2

- 5.3d. *Our analysis.* The images of the gels provided by Respondents for p62 and Nrf2 have significantly different exposures from the ones used in the published figure. The data presented in the published figure are consistent with the original data.
- 5.3e. *Relevant interview questions/comments.* Not applicable.
- 5.3f. *Conclusion for this specific example.* The Respondents did not provide the specific data used to generate the published figure. This is a significant departure from accepted

practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

6. Incomplete metadata provided with figure components:

Description: The committee defined this category as images and data that lacked experimental information such as dates, exposure times, magnifications, etc.

Overall summary: 6 figures from grants and 7 figure panels from manuscripts. (sections 6.1-6.13).

Overall summary of interviews and relevant comments: The committee did not ask specific questions about this category.

Overall conclusions: This category contains fluorescence imaging data where comparisons between treatment groups are an essential part of the experiment. The Respondents rely on image brightness as a metric of the signal intensity. However, if the exposure times are not the same, the comparisons are not valid. The data provided by the Respondents were lacking metadata that indicated the exposure times and other imaging parameters. Therefore, the committee was unable to determine whether similar exposure times were used in a given figure and thus whether the images could be validly compared.

Based on the committee's examination of a number of images provided by the Respondents, it is unclear how imaging metadata is saved when images are taken. Some images had complete sets of metadata (camera type and model, exposure time, etc.) while others taken with the same camera systems did not. Thus, the deficits noted in this category could be operator error; however, if that were the case, it reflects a lack of appropriate training on the microscope systems. Where the images were obviously cropped or otherwise manipulated, the committee decided that the research misconduct was intentional. Where the metadata was omitted from what appeared to be original images, the committee considered the conduct to be reckless.

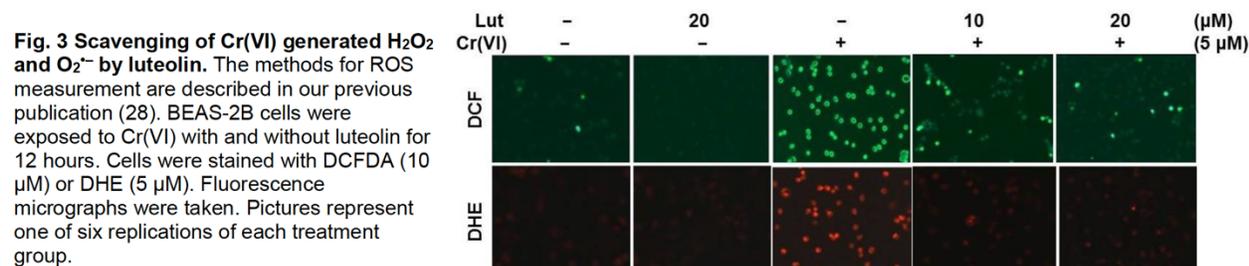
Specific examples for "incomplete metadata provided with figure components":

6.1 G2: Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Fig. 3

6.1a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

6.1b. *Original figure from Appendix 007 page 2 of the research strategy section of the grant.*



- 6.1c. *Data provided.* See Appendix 122. The Respondents provided the following statement along with the data: “This work (Fig 3) was done by Dr. Poyil Pratheeskumar. ... Dr. Pratheeskumar provided the results in PPT for Fig. 3 in the grant”.
- 6.1d. *Our analysis.* Only JPG files of the composite DCF and DHE images were provided. The individual panels were not provided, and therefore, the committee could not determine imaging parameters such as exposure times, image sizes, magnification, etc. Thus, the committee could not determine if the images were correctly compared in the figure.
- 6.1e. *Relevant interview questions/comments.* Not applicable.
- 6.1f. *Conclusion for this specific example.* The raw imaging files should have been stored in the laboratory. Without these files, the committee cannot assess the veracity of the figure and thus has to conclude that it was fabricated. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence.
- 6.2. G2: Grant Shi_3048112536 = 1R01ES025515-01 (05/01/2015-01/31/2020) (Appendix 007): Fig. 13
- 6.2a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 6.2b. *Original figure from Appendix 007 page 13 of the research strategy section of the grant.*

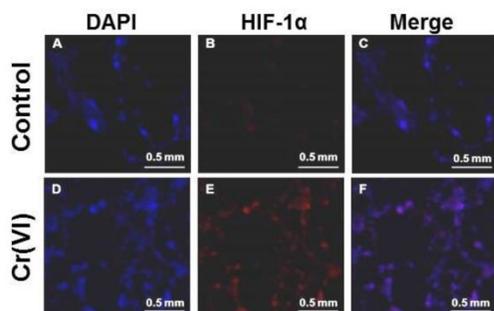


Fig.13 Increase of HIF-1 α expression in the lung tissue of animal exposed to Cr(VI) via drinking water. 6-8 weeks, female C57B/6 mice were randomly separated into different treatment group with 10 mice/group. The mice were treated with or without 167 mg/L of sodium dichromate dehydrate in fresh drinking water. After 20 weeks, the mice were euthanized by CO₂ and lung tissues were isolated. Tissue sections of lung were used to examine the expression of HIF-1 α . A and D, fluorescence staining of DAPI in control and Cr(VI) treatment groups. B and E, fluorescence staining of HIF-1 α in control and Cr(VI) treatment groups. C and F, co-localization of DAPI and HIF-1 α . Sections are representative of 10 different mice from each treatment group.

- 6.2c. *Data provided.* See Appendix 123.
- 6.2d. *Our analysis.* Images for all 6 panels were provided as TIF files. None of these images had scale bars embedded. Panels d-f had appropriate metadata (camera system, exposure time and image dimensions in pixels). Panels a-c were provided as images modified with Photoshop and lacked metadata. Based on the image dimensions, panels a-c were cropped

and appeared to be at a different magnification in the figure (based on the size of the DAPI-stained nuclei). Based on the images and lack of embedded scale bars, it is not possible to confirm the accuracy of the scale bars included in the grant figure.

6.2e. *Relevant interview questions/comments.* Not applicable.

6.2f. *Conclusion for this specific example.* Without the metadata and embedded scale bars, the accuracy of the images presented cannot be confirmed and thus the committee has to conclude that the data used in the grant were fabricated. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence.

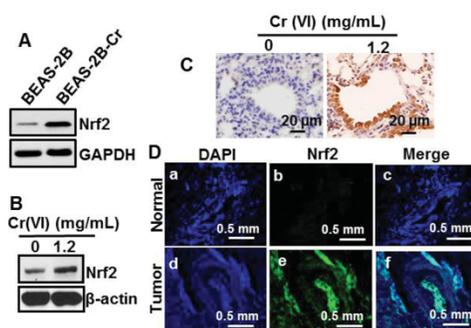
6.3. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008): Fig. 7D

6.3a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

6.3b. *Original figure from Appendix 008 page 54.*

Fig. 7 Increased Nrf2 protein level in Cr(VI)-transformed cells, in lung tissues from Cr(VI)-exposed animals, and in lung tumor tissues from a worker exposed to Cr(VI) for 19 years. (A) Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) were cultured in 10-cm dishes. Whole protein was isolated for examination of Nrf2 level using immunoblotting. The results are representative of three independent experiments. (B) and (C) 4-week old both male and female BALB/c mice were intranasally exposed to Cr(VI) particles (1.2 mg/mL) as described in Fig. 2. Lung tissues were isolated to examine Nrf2 level using immunoblotting (B) and immunohistological (C) analyses. Intensity of brown color represents Nrf2 level (C). Sections are representative of 6 different animals from each treatment group. (D) Formalin-fixed human lung tumor and adjacent normal tissues from a non-smoking worker exposed to Cr(VI) were subjected to fluorescence immunostainings with DAPI (nuclear control, a and d) and Nrf2 (b and e). The merged images are provided in c and f. The results are representative of three independent experiments.



6.3c. *Data provided.* See Appendix 116.

6.3d. *Our analysis.* Images for all 6 panels were provided as JPG files. Only panel c had appropriate metadata (exposure time). The scale bar issues in this figure have been discussed in a previous section.

6.3e. *Relevant interview questions/comments.* Not applicable.

6.3f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing panels b and e cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.

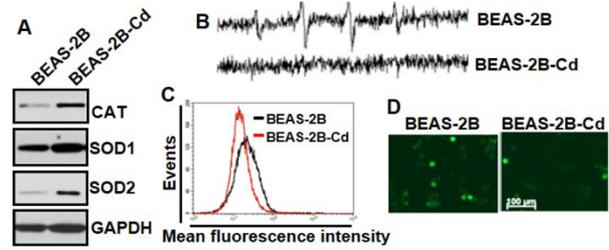
6.4. G6: Grant Zhang_3200001638 = 1R01ES028984-01 (12/15/2017-11/30/2022) (Appendix 011): Fig. 8D

6.4a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents via flash drive.

6.4b. Original figure from Appendix 011 page 5 of the research strategy section of the grant.

Fig. 8 Increased expressions of antioxidant enzymes and reduced ROS levels in Cd(II)-transformed cells. (A) Levels of catalase (CAT), SOD1, and SOD2 were examined in passage-matched normal cells (BEAS-2B) and in Cd(II)-transformed cells (BEAS-2B-Cd) by immunoblotting. (B) The cell suspensions were prepared from normal cells (BEAS-2B) or Cd(II)-transformed cells (BEAS-2B-Cd). ESR spectra representing ROS generation were recorded. ROS levels in passage-matched normal cells (BEAS-2B) and in Cd(II)-transformed cells (BEAS-2B-Cd) were measured by flow cytometry (C) and fluorescence microscopy (D). The results in each subfigure are representative of three independent experiments.



6.4c. Data provided. See Appendix 124.

6.4d. Our analysis. The only metadata contained in the TIF files provided is the camera model, with no information on the exposure times or other imaging parameters.

6.4e. Relevant interview questions/comments. Not applicable.

6.4f. Conclusion for this specific example. Without recorded exposure times, the validity of comparing the two panels cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.

6.5. G6: Grant Zhang_3200001638 = 1R01ES028984-01 (12/15/2017-11/30/2022) (Appendix 011): Fig. 11

6.5a. Date requested: October 18, 2018 by the committee via email.

Date received: October 26, 2018 by Respondents via flash drive.

6.5b. Original figure from Appendix 011 page 6 of the research strategy section of the grant.

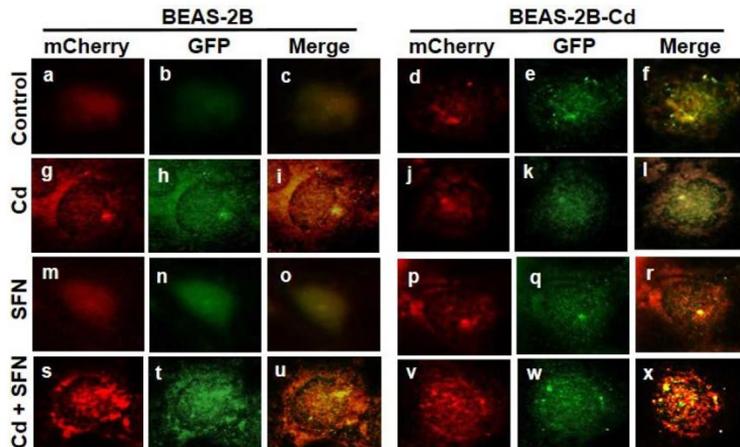


Fig.11 Sulforaphane increases Cd(II)-induced autophagy in normal and Cd(II)-transformed cells. Passage-matched normal cells (BEAS-2B) (left) and Cd(II)-transformed cells (BEAS-2B-Cd) (right) were transfected with the mCherry-EGFP-LC3 construct as described in Fig. 4. The cells were treated with Cd(II), sulforaphane (SFN), or Cd(II) together with SFN for 24 h. Images were captured using fluorescence microscope. The images represent one of three independent samples in each group.

6.5c. *Data provided.* See Appendix 125.

6.5d. *Our analysis.* The Respondents provided PNG files for each panel in the figure. These images were cropped and had different sizes. No metadata were contained in these files.

6.5e. *Relevant interview questions/comments.* Not applicable.

6.5f. *Conclusion for this specific example.* Since the images were different sizes and contained no metadata about exposure times, the validity of comparing the many panels cannot be confirmed and thus the committee has to conclude that the images were fabricated.. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence.

6.6. G7: Grant Zhang_3200001897 = 1R01CA228236-01A1 (01/01/2018-05/31-2023) (Appendix 012): Fig. 9

6.6a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 24, 2018 by Respondents *via* flash drive.

6.6b. *Original figure from Appendix 012 page 64 of the research strategy section from the grant.*

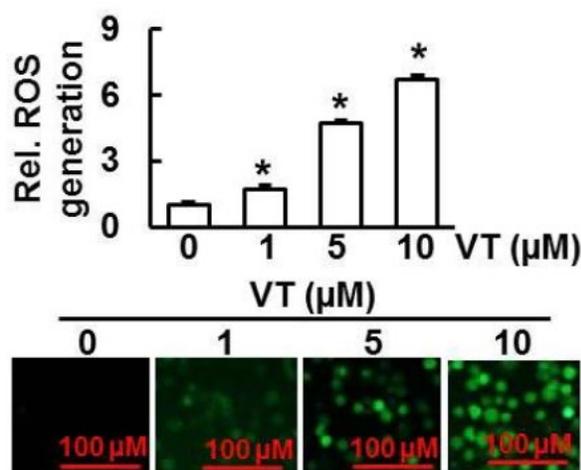


Fig 9 Verteporfin increases ROS generation in PC3 cells. PC3 cells were treated with various doses of verteporfin for 6 h followed by staining with 10 μM 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 minutes. Fluorescence intensity was measured by flow cytometry (Upper). The results are shown as mean±SE (n=6). Images were taken by a fluorescence microscope (Bottom). The images represent one of six samples from each treatment group. *, $p < 0.05$ compared to control without treatment.

6.6c. *Data provided.* See Appendix 123.

- 6.6d. *Our analysis.* The only metadata contained in the TIF files provided is the camera model, with no information on the exposure times or other imaging parameters. There were no scale bars on the provided images, and the scale bars in the grant figure have the incorrect units (another example of problems noted in section 4).
- 6.6e. *Relevant interview questions/comments.* Not applicable.
- 6.6f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing the four panels cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.
- 6.7. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 3B
- 6.7a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 6.7b. *Original figure from Appendix 013, page 395.*

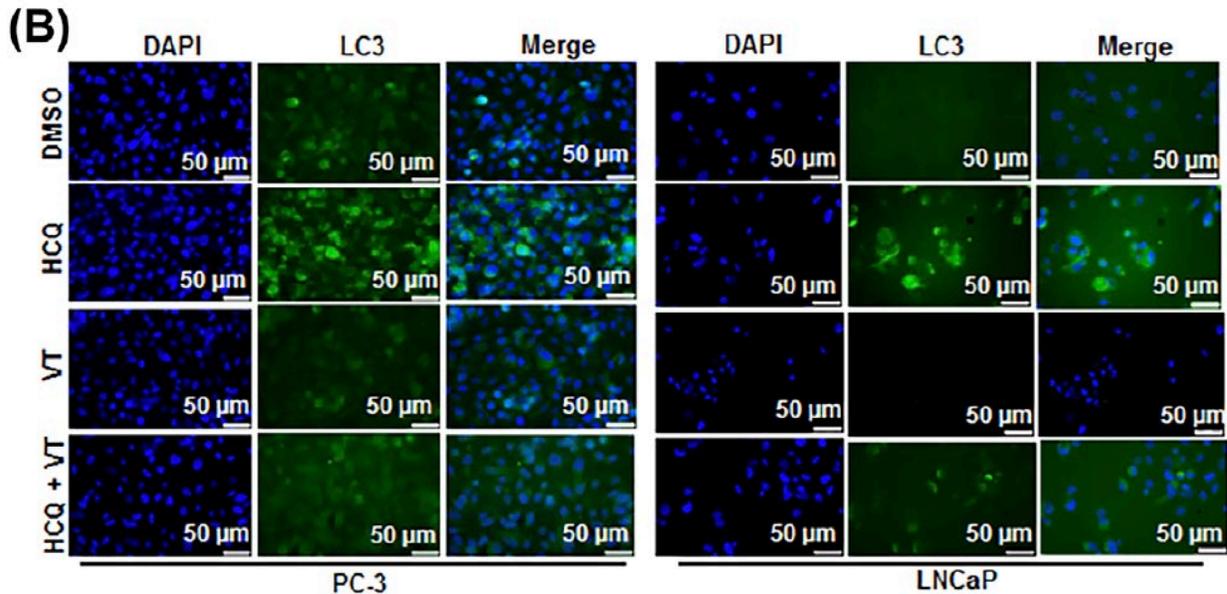


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μ M Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μ M and 10 μ M Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μ M 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean \pm SE ($n = 6$). * and #, $P < 0.05$ compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μ M of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

6.7c. *Data provided.* See Appendix 127.

6.7d. *Our analysis.* The data provided for Fig. 3B do match the published figure. However, the TIF files provided contained only the camera model with no information on the exposure time. Without the exposure times for the green channel, it is hard to determine the validity of comparing these images. It should be noted that when the committee used Photoshop they could find cells in all the green images.

6.7e. *Relevant interview questions/comments.* Not applicable.

6.7f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing the GFP-LC3 panels cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.

6.8. M5: Wang, L., Son, Y.O., Ding, S., Wang, X., Hitron, J.A., Budhraj, A., Lee, J.C., Lin, Q., Poyil, P., Zhang, Z., Luo, J., and Shi, X. Ethanol enhances tumor angiogenesis in vitro induced by low-dose arsenic in colon cancer cells through hypoxia-inducible factor 1 alpha pathway. *Toxicological Sciences* 130, 269-280, 2012 (Appendix 017): Fig. 2C

6.8a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

6.8b. *Original figure from Appendix 017, page 273.*

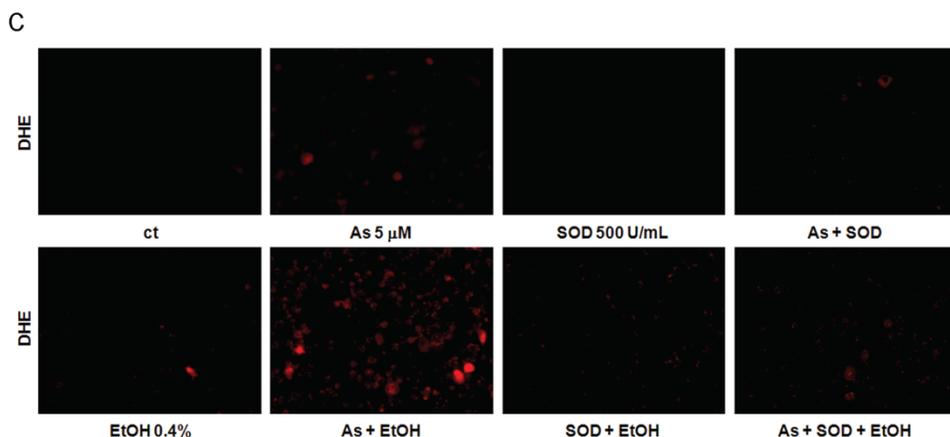


FIG. 2. Low-dose arsenic combined with ethanol induces ROS generation in colon cancer cells. (A) DLD-1 cells were exposed to arsenic and/or ethanol at indicated concentrations for 24 h and then stained with 10 μ M H₂DCFDA or 5 μ M DHE, respectively for 30 min. Cells were imaged by fluorescence microscopy. (B) Cells were incubated with H₂DCFDA or oxidized DCFDA. The fluorescence in cells was measured with a fluorescence microplate reader. The ratios of H₂DCFDA/DCFDA were calculated between different exposure groups. (C) DLD-1 cells were exposed to arsenic and/or ethanol at the indicated concentrations with or without 500U/ml SOD for 24h and then stained with 5 μ M DHE. Cells were imaged by fluorescence microscopy. (D) Cells were exposed to arsenic and/or ethanol with and without 500 U/ml catalase, then incubated with H₂DCFDA or oxidized DCFDA. The fluorescence in cells was measured with fluorescence microplate reader. Columns, mean of six duplicates; bars, SE (* $p \leq 0.05$).

- 6.8c. *Data provided.* See Appendix 128.
- 6.8d. *Our analysis.* The only metadata contained in the TIF files provided is the camera model, with no information on the exposure times or other imaging parameters.
- 6.8e. *Relevant interview questions/comments.* Not applicable.
- 6.8f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing the different panels cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.
- 6.9. M9: Kim, D., Dai, J., Park, Y.H., Yenwong F., L., Wang, L., Pratheeshkumar, P., Son, Y.O., Kondo, K., Xu, M., Luo, J., Shi, X., and Zhang, Z. Activation of EGFR/p38/HIF-1 α is pivotal for angiogenesis and tumorigenesis of malignantly transformed cells induced by hexavalent chromium. *Journal of Biological Chemistry* 291, 16271-16281, 2016 (Appendix 021): Fig. 2Hb
- 6.9a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 6.9b. *Original figure from Appendix 021, page 16273.*

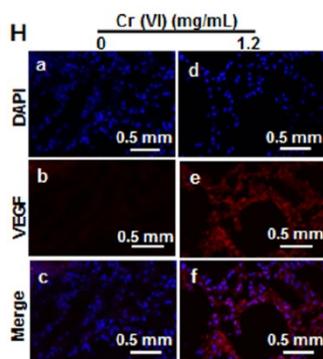


FIGURE 2. Increased angiogenic proteins in Cr(VI)-transformed cells. *A*, angiogenesis protein array. *Box 1*, GM-CSF; *box 2*, MMP-1; *box 3*, angiogenin; *box 4*, IL-6; *box 5*, TGF- α ; *box 6*, MCP-1. *B*, quantitative PCR analysis. mRNA levels of angioenin, IL-6, MMP-1, and GM-CSF were measured using real time PCR in Cr(VI)-transformed cells (BEAS-2B-Cr) and their passage-matched normal ones (BEAS-2B). *, $p < 0.05$, statistically significant difference compared with BEAS-2B cells. *C* and *E*, whole cell lysates were collected in BEAS-2B-Cr and BEAS-2B cells and subjected to immunoblotting analysis. *D*, production of VEGF was measured using ELISA. *, $p < 0.05$, compared with BEAS-2B cells. *F–H*, formalin-fixed lung tissue from animals exposed to Cr(VI) or saline was used for immunofluorescence staining analysis. *I* and *J*, formalin-fixed normal lung tissue from a healthy patient and lung tumor tissue and its adjacent normal tissue from a worker diagnosed with stage 1 lung adenocarcinoma caused by occupational exposure to Cr(VI) were subjected to immunofluorescence staining for examination of expression of MMP-1 (*I*), HIF-1 α (*J*, green), and VEGF (*J*, red). DAPI (blue) was used for nuclear control.

- 6.9c. *Data provided.* See Appendix 129. The data provided was a PDF of the figure with some selected panels shown with enhanced brightness. The Respondents also provided the file paths for the original data. Following those file paths, the committee found the relevant image files and, using Photoshop, demonstrated that they contained data.
- 6.9d. *Our analysis.* Only limited metadata were associated with the JPG files (camera system, date, metering mode and subject distance), but no exposure times.
- 6.9e. *Relevant interview questions/comments.* Not applicable.
- 6.9f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing protein expression in panels b and e cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.
- 6.10. M9: Kim, D., Dai, J., Park, Y.H., Yenwong F., L., Wang, L., Pratheeshkumar, P., Son, Y.O., Kondo, K., Xu, M., Luo, J., Shi, X., and Zhang, Z. Activation of EGFR/p38/HIF-1 α is pivotal for angiogenesis and tumorigenesis of malignantly transformed cells induced by hexavalent chromium. *Journal of Biological Chemistry* 291, 16271-16281, 2016 (Appendix 021): Fig. 2Ib
- 6.10a. *Date requested:* October 24, 2018 by the committee via email.
Date received: November 2, 2018 by Respondents via flash drive.
- 6.10b. *Original figure from Appendix 021, page 16273.*

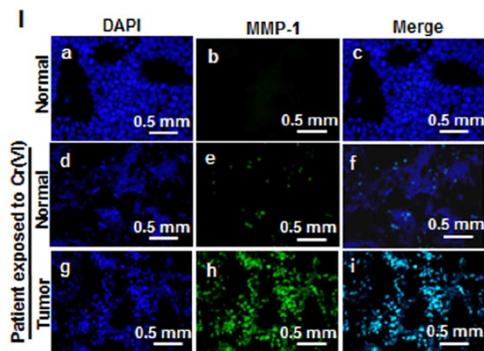


FIGURE 2. **Increased angiogenic proteins in Cr(VI)-transformed cells.** A, angiogenesis protein array. Box 1, GM-CSF; box 2, MMP-1; box 3, angiogenin; box 4, IL-6; box 5, TGF- α ; box 6, MCP-1. B, quantitative PCR analysis. mRNA levels of angiogenin, IL-6, MMP-1, and GM-CSF were measured using real time PCR in Cr(VI)-transformed cells (BEAS-2B-Cr) and their passage-matched normal ones (BEAS-2B). *, $p < 0.05$, statistically significant difference compared with BEAS-2B cells. C and E, whole cell lysates were collected in BEAS-2B-Cr and BEAS-2B cells and subjected to immunoblotting analysis. D, production of VEGF was measured using ELISA. *, $p < 0.05$, compared with BEAS-2B cells. F–H, formalin-fixed lung tissue from animals exposed to Cr(VI) or saline was used for immunofluorescence staining analysis. I and J, formalin-fixed normal lung tissue from a healthy patient and lung tumor tissue and its adjacent normal tissue from a worker diagnosed with stage 1 lung adenocarcinoma caused by occupational exposure to Cr(VI) were subjected to immunofluorescence staining for examination of expression of MMP-1 (I), HIF-1 α (J, green), and VEGF (J, red). DAPI (blue) was used for nuclear control.

- 6.10c. *Data provided.* See Appendix 130. The data provided was a PDF of the figure with some selected panels shown with enhanced brightness. The Respondents also provided the file paths for the original data. Following those file paths, the committee found the relevant image files and, using Photoshop, demonstrated that they contained data.
- 6.10d. *Our analysis.* Only limited metadata were associated with the JPG files (camera system, date, metering mode and subject distance), but no exposure times.
- 6.10e. *Relevant interview questions/comments.* Not applicable.
- 6.10f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing protein expression in panels b, e, and h cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.
- 6.11. M9: Kim, D., Dai, J., Park, Y.H., Yenwong F., L., Wang, L., Pratheeshkumar, P., Son, Y.O., Kondo, K., Xu, M., Luo, J., Shi, X., and Zhang, Z. Activation of EGFR/p38/HIF-1 α is pivotal for angiogenesis and tumorigenesis of malignantly transformed cells induced by hexavalent chromium. *Journal of Biological Chemistry* 291, 16271-16281, 2016 (Appendix 021): Fig. 6Bb
- 6.11a. *Date requested:* October 24, 2018 by the committee via email.
Date received: November 2, 2018 by Respondents via flash drive.
- 6.11b. *Original figure from Appendix 021, page 16276.*

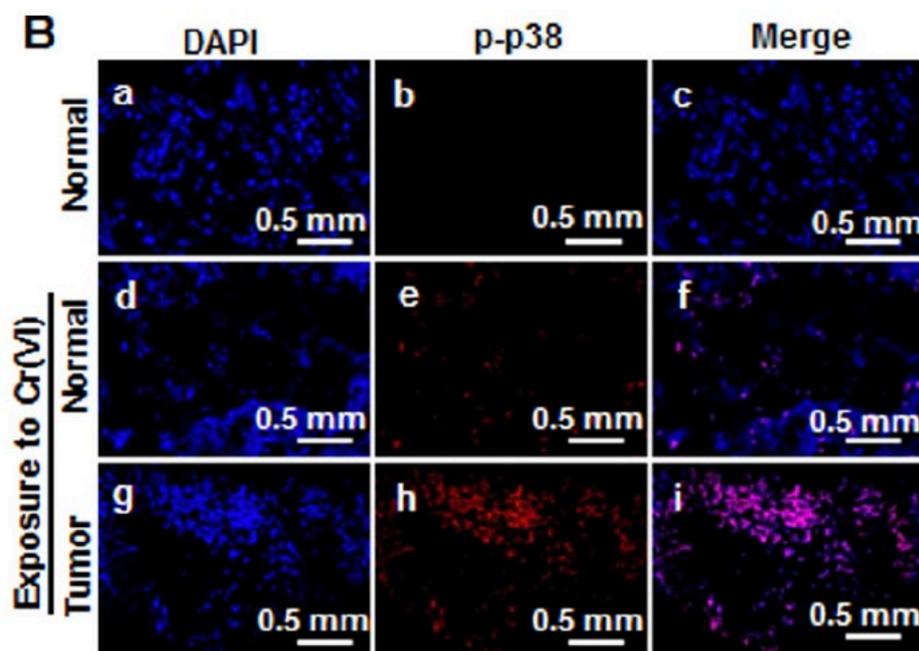


FIGURE 6. Inhibition of EGFR down-regulates HIF-1 α and p38 in Cr(VI)-transformed cells. *A*, Cr(VI)-transformed cells (BEAS-2B-Cr) and passage-matched normal ones (BEAS-2B) were collected for immunoblotting analysis. *B*, immunofluorescence staining analysis in normal lung tissue from a health patient and in lung tumor tissue and its normal adjacency from a worker exposed to Cr(VI) for 19 years as described under “Experimental Procedures.” *C*, binding of HIF-1 α to hypoxia response element region of VEGF promoter was examined using ChIP assay. Data are means \pm S.D. ($n = 3$). * and #, $p < 0.05$, compared with BEAS-2B cells and BEAS-2B-Cr scramble cells. *D*, BEAS-2B-Cr cells with (shEGFR) and without (Scramble) (*left panel*) and tumor tissues injected with stable knockdown EGFR (shEGFR) and scramble (*right panel*) were subjected to immunoblotting analysis. *E*, BEAS-2B-Cr cells were treated with SB203580 or PX478 for 24 h. Whole cell lysates were harvested for immunoblotting analysis. *F*, tube formation assay. Conditioned medium were collected from BEAS-2B-Cr cells treated with SB203580 or PX478 for 24 h or without treatment. HUVECs were plated onto Matrigel-coated plates and incubated with those conditioned medium. *Top panel*, representative images from each treatment group. *Bottom panel*, quantitative results of tube formation are shown as mean \pm S.D. ($n = 3$). *, $p < 0.05$, compared with BEAS-2B-Cr cells without treatment. *G*, BEAS-2B-Cr cells were transiently transfected with HIF-1 α shRNA for 48 h. Whole cell lysates were harvested for immunoblotting analysis. *H*, Cr(VI)-transformed cells were transfected with EGFR expression plasmid and/or EGFR shRNA for 48 h in various combinations. The cells were collected, and whole cell lysates were subjected for immunoblotting analysis.

- 6.11c. *Data provided.* See Appendix 131. The data provided was a PDF of the figure with some selected panels shown with enhanced brightness. The Respondents also provided the file paths for the original data. Following those file paths, the committee found the relevant image files and, using Photoshop, demonstrated that they contained data.
- 6.11d. *Our analysis.* Only limited metadata were associated with the JPG files (camera system, date, metering mode and subject distance), but no exposure times.
- 6.11e. *Relevant interview questions/comments.* Not applicable.
- 6.11f. *Conclusion for this specific example.* Without recorded exposure times, the validity of comparing protein expression in panels b, e, and h cannot be confirmed and thus the committee has to conclude that the images were falsified. This is a significant departure from accepted practices of the research community that was committed recklessly and this allegation is proven by a preponderance of evidence.
- 6.12. M12: Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROS-dependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Fig. 2D

- 6.12a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 6.12b. *Original figure from Appendix 024, page 4.*

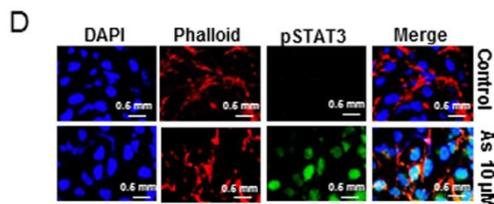


Figure 2. Arsenic initiates an IL-6-mediated STAT3 binding to the miR-21 promoter. (A–C) BEAS-2B cells were treated with arsenic (5 and 10 μ M) for 24 h. (A) Increased IL-6 levels in culture medium and (B) STAT3 activity in the nuclear fraction of cell lysate were assessed using ELISA kits, following the manufacturer’s protocol. H₂O₂ (0.1 mM) and IL-6 (10 ng/ml) were used as positive control. (C) Total cell lysates were prepared for western blot analysis using specific antibodies against pSTAT3 and STAT3. Arsenic induced an apparent increase in activated levels in the nucleus. (D) Representative images of fluorescence immunostaining confirm presence of nuclear pSTAT3 with arsenic treatment (10 μ M). Dapi: blue-nuclear; Phalloid: red-cytoplasmic actin; pSTAT3, green (E) Reporter gene assays were performed in BEAS-2B cells transfected with either a luciferase vector driven by the miR-21 promoter/enhancer alone (–) or in the presence of a vector encoding a small hairpin RNA silencing STAT3 expression (siSTAT3). Knockdown of STAT3 reduced promoter activity. (F–G) BEAS-2B cells were exposed to IL-6 or arsenic for 30 minutes and subjected to a chromatin immunoprecipitation (ChIP) analysis using anti-STAT3 or IgG isotype control. Co-immunoprecipitated DNA was amplified by PCR with primers specific for the miR-21 upstream enhancer. Both arsenic and IL-6 increased binding to the mi-R21 promoter. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *Indicates a statistically significant difference compared to control with $p < 0.05$.

- 6.12c. *Data provided.* See Appendix 132. The Respondents provided only a PDF file that contained the panels from the published figure. No original images or file paths to original images were provided.
- 6.12d. *Our analysis.* Without image files, there is no information on the imaging parameters that could be analyzed by the committee.
- 6.12e. *Relevant interview questions/comments.* Not applicable.
- 6.12f. *Conclusion for this specific example.* Without the metadata, the validity of the comparisons between the different panels in this figure cannot be validated and thus the committee has to conclude that the images were fabricated. It should be noted that this figure also shows up in section 7.2. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence.
- 6.13. M12: Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROS-dependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Fig. 3E
- 6.13a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 6.13b. *Original figure from Appendix 024, page 5.*

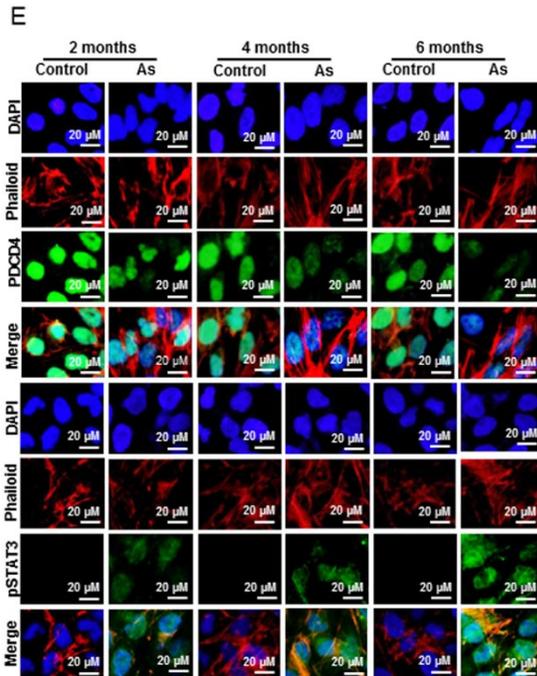


Figure 3. The arsenic-induced miR-21 increase and PDCD4 suppression contribute to malignant cell transformation. BEAS-2B cells were maintained in a medium containing various concentrations of arsenic (0.1, 0.25 and 0.5 μ M) for 6 months. (A,B) Cells were cultured in 0.35% soft agar for 5 weeks and number of colonies in the entire dish counted. (A) Representative images of control (left panel) and arsenic-treated (right panel) colonies. (B) Colony number increased in a dose-dependent manner. (C) The relative miR-21 level, determined by Taqman real-time PCR, increased in a time- and dose-dependent manner. (D) Total cell lysates were prepared for western blot analysis after 2, 4 and 6 months exposure to arsenic using specific antibodies against PDCD4, p47phox, pSTAT3 and STAT3. Apparent protein levels for PDCD4 decreased and P47phox and pSTAT3 increased in a time- and dose-dependent manner. (E) Representative images of fluorescence immunostaining for PDCD4 and pSTAT3 after 2, 4 and 6 months exposing to arsenic, and confirm results from western blot analysis. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *Indicates a statistically significant difference compared to control with $p < 0.05$.

- 6.13c. *Data provided.* See Appendix 119. The Respondents provided only a PDF file that contained the panels from the published figure. No original images or file paths to original images were provided.
- 6.13d. *Our analysis.* Without image files, there is no information on the imaging parameters that could be analyzed by the committee.
- 6.13e. *Relevant interview questions/comments.* The committee only asked questions about the scale bars in these figures and did not address the absence of metadata.
- 6.13f. *Conclusion for this specific example.* Without the metadata, the validity of the comparisons between the different panels in this figure cannot be validated and thus the committee has to conclude that the images were fabricated. It should be noted that this figure shows up in two other categories (sections 4.5 and 7.3). This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence.

7. Black images containing no pixel data:

Description: The committee defined this category as images where, upon analysis with Photoshop, no signal was detected.

Overall summary: 1 figure from grants and 2 figures from manuscripts. (sections 7.1-7.3).

Overall summary of interviews and relevant comments: The committee only asked questions about one figure, and the responses from Respondents are included in section 7.1 below.

Overall conclusions: In the three figures that the committee examined in detail, using Photoshop tools (below), there were black panels that had no pixel data, indicating no fluorescence signal. Given the general sensitivity of camera systems, generating a completely black image from an immunofluorescence experiment is highly unlikely. The committee expected to at least see evidence of background fluorescence in these images, but there was none. During a visit to the lab on October 19, 2018 (Appendix 035, DVD), the committee witnessed Dr. Kim take images of blank slides. Upon analysis, the committee found that these images contained no pixel data, similar to the panels highlighted in this section. The committee concludes that the images included in these figures did not represent real experimental data but were either of blank slides or were generated using a graphics program (*i.e.*, PowerPoint).

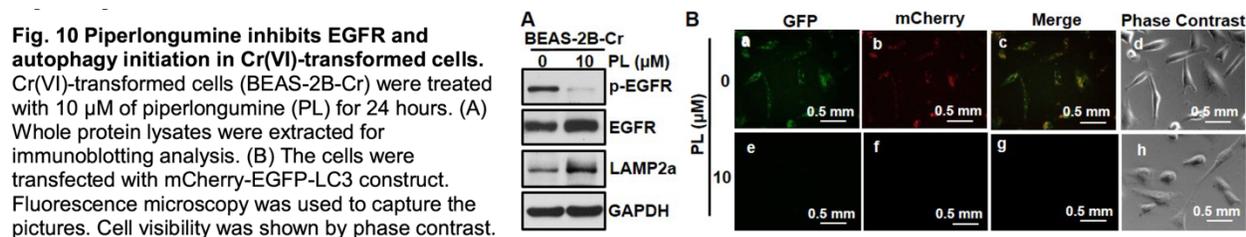
It should be noted that the Federal Office of Research Integrity highlighted one of these figures (section 7.3 below) as being suspect of having no pixel data.

Specific examples for “black images containing no pixel data”:

7.1 G1: Grant Shi_3210000529 = 3R01ES025515-03S1 (02/01/2017-01/30/2020) (Appendix 006, pages 11-133): Fig. 10Bg

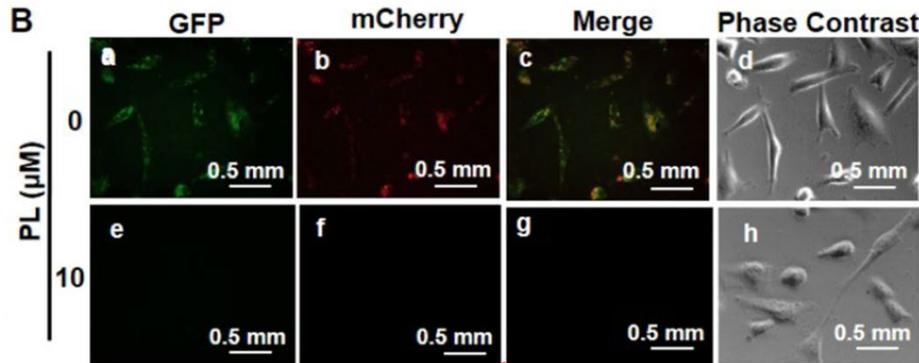
7.1a. *This figure was originally provided to the committee at the start of the investigation.*

7.1b. *Original figure from Appendix 006 page 63 of the research strategy section of the grant..*



7.1c. *Data provided.* This figure was originally provided to the investigation committee in the report of the Inquiry Committee (Appendix 006).

7.1d. *Our analysis.* The committee notes that image 10Bg was part of the original complaint. The committee analyzed the images of the figure extracted from the PDF version of the grant using the Photoshop Levels Tool. Consistent with the original complaint and the inquiry committee finding, our analysis showed that there was no pixel data in the image. During our observations of Dr. Kim taking images with their microscope system, the committee had him take images of a blank slide, with no illumination. Upon analysis with Photoshop, these image files had no data in them, similar to the image in Fig. 10Bg.



FROM:
Grant ES025515
Supplement

Figure 10B-g

Fig. 10 Piperlongumine inhibits EGFR and autophagy initiation in Cr(VI)-transformed cells. Cr(VI)-transformed cells (BEAS-2B-Cr) were treated with 10 μ M of piperlongumine (PL) for 24 hours. (A) Whole protein lysates were extracted for immunoblotting analysis. (B) The cells were transfected with mCherry-EGFP-LC3 construct. Fluorescence microscopy was used to capture the pictures. Cell visibility was shown by phase contrast.



- 7.1e. *Relevant interview questions/comments.* Dr. Shi stated that he cannot explain why there is no data in this figure. The data were provided by Dr. Kim to Dr. Zhang, and Dr. Zhang inserted it directly into the grant. Dr. Shi stated that he was not involved (Appendix 052 (Shi interview, page 41)). Dr. Shi stated that he did not see the raw data when he was assembling the grant and did not check the raw data (Appendix 052 (Shi interview, page 42)). Dr. Zhang confirmed that Dr. Kim took the images; however, she does not know why there is no data in panel 10Bg (Appendix 054 (Zhang interview, pages 43-44)). Dr. Kim performed the experiment once and showed it at a lab meeting (Appendices 054 (Zhang interview, pages 43-44) and 063 (Kim interview II, page 13)). The image appeared blank, and Dr. Zhang and another lab member asked Dr. Kim if the cells were still alive (Appendix 063 (Kim interview II, page 13)). Dr. Zhang asked Dr. Kim to repeat the experiment (Appendix 063 (Kim interview II, page 14)). Dr. Zhang confirmed that Dr. Kim did the experiment a second time (Appendix 054 (Zhang interview, pages 43-44)). Dr. Kim took a phase contrast image to show that the cells were still alive, and he imaged six fields on the slide (Appendix 063 (Kim interview II, page 14)). All six fields were blank with the same exposure and contrast. The merged image was created by the computer software (Appendix 063 (Kim interview II, page 16)). Dr. Kim stated that he used Fig. 6 (from the same grant, which was also part of the original complaint) as a template and overlaid Fig. 10 on Fig. 6. The bottom slide on both Figures was blank (Appendix 063 (Kim interview II, page 15)).
- 7.1f. *Conclusion for this specific example.* Fig. 10Bg did not show any evidence of fluorescence signal or background noise, and thus had no data in it. Using a black box instead of actual data is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 7.2 M12: Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROSdependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Fig. 2D
- 7.2a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 7.2b. *Original figure from Appendix 024, page 4.*

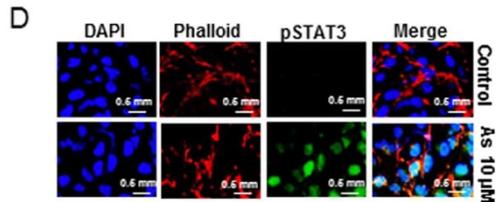
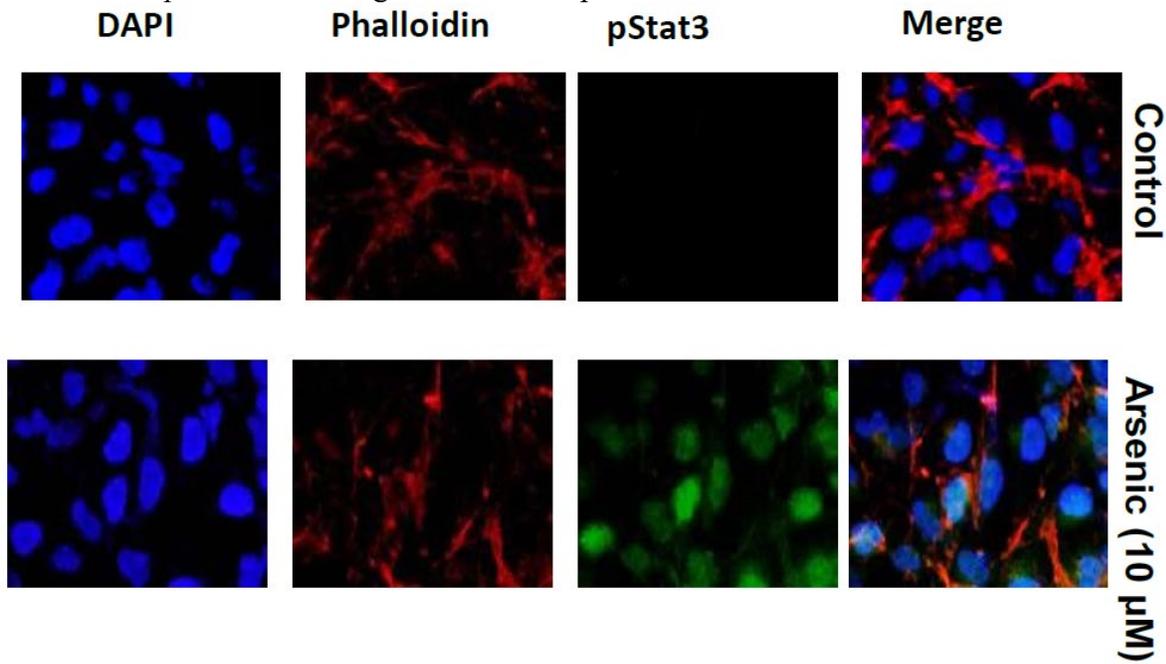
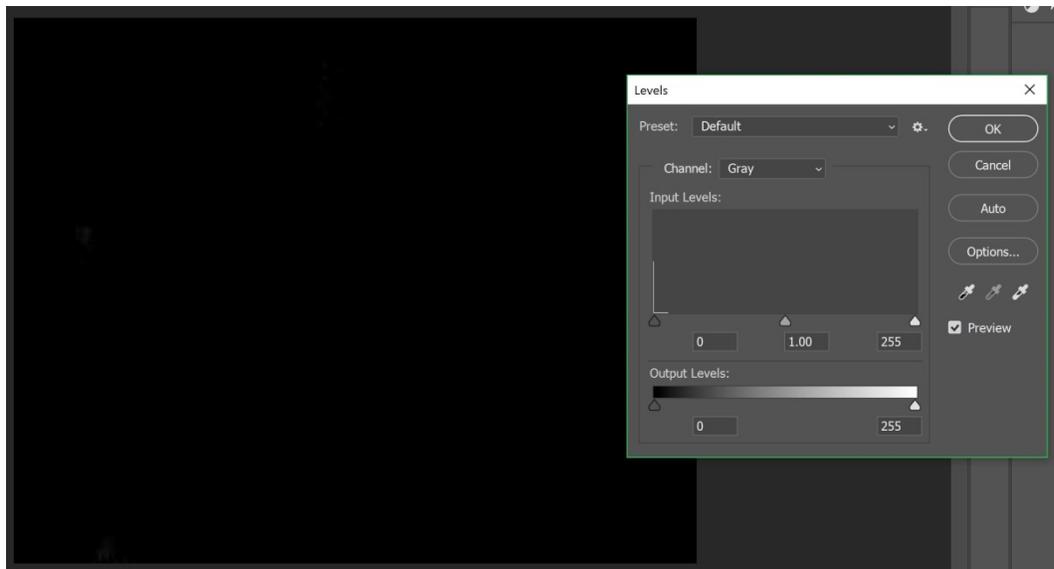


Figure 2. Arsenic initiates an IL-6-mediated STAT3 binding to the miR-21 promoter. (A–C) BEAS-2B cells were treated with arsenic (5 and 10 μ M) for 24 h. (A) Increased IL-6 levels in culture medium and (B) STAT3 activity in the nuclear fraction of cell lysate were assessed using ELISA kits, following the manufacturer’s protocol. H₂O₂ (0.1 mM) and IL-6 (10 ng/ml) were used as positive control. (C) Total cell lysates were prepared for western blot analysis using specific antibodies against pSTAT3 and STAT3. Arsenic induced an apparent increase in activated levels in the nucleus. (D) Representative images of fluorescence immunostaining confirm presence of nuclear pSTAT3 with arsenic treatment (10 μ M). Dapi: blue-nuclear; Phalloidin: red-cytoplasmic actin; pSTAT3, green (E) Reporter gene assays were performed in BEAS-2B cells transfected with either a luciferase vector driven by the miR-21 promoter/enhancer alone (–) or in the presence of a vector encoding a small hairpin RNA silencing STAT3 expression (siSTAT3). Knockdown of STAT3 reduced promoter activity. (F–G) BEAS-2B cells were exposed to IL-6 or arsenic for 30 minutes and subjected to a chromatin immunoprecipitation (ChIP) analysis using anti-STAT3 or IgG isotype control. Co-immunoprecipitated DNA was amplified by PCR with primers specific for the miR-21 upstream enhancer. Both arsenic and IL-6 increased binding to the mi-R21 promoter. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *Indicates a statistically significant difference compared to control with $p < 0.05$.

- 7.2c. *Data provided.* The figure below was provided in a PDF file.



- 7.2d. *Our analysis.* In our analysis, the panel for pSTAT3 Control in the published Fig. 2D did not have any pixel data in it. The image used for the panel was black. A caveat to our analysis is that the committee had to work with a PDF of the original data rather than the original image file itself, as the Respondents did not provide the raw images. The Respondents included the following statement with the provided data: “Dr. Pratheeshkumar Poyil states that there was no basal expression of phosphor Sta3 found in Beas-2B cell by IF. Unstained cells cannot be detected even after increase brightness.” (Appendix 119)



- 7.2e. *Relevant interview questions/comments.* Not applicable.
- 7.2f. *Conclusion for this specific example.* The black panel in Fig. 2D did not show any evidence of fluorescence signal or background noise, and thus had no data in it. This is identical to the images of blank slides that Dr. Kim took during the committee’s visit to the lab. Using a black box instead of actual data is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.
- 7.3 M12: Pratheeshkumar, P., Son, Y.O., Divya, S.P., Wang, L., Zhang, Z., and Shi X. Oncogenic transformation of human lung bronchial epithelial cells induced by arsenic involves ROSdependent activation of STAT3-miR-21-PDCD4 mechanism. *Scientific Reports* 6, 37227, 2016 (Appendix 024): Fig. 3E
- 7.3a. *Date requested:* October 24, 2018 by the committee *via* email.
Date received: November 2, 2018 by Respondents *via* flash drive.
- 7.3b. *Original figure from Appendix 024, page 5.*

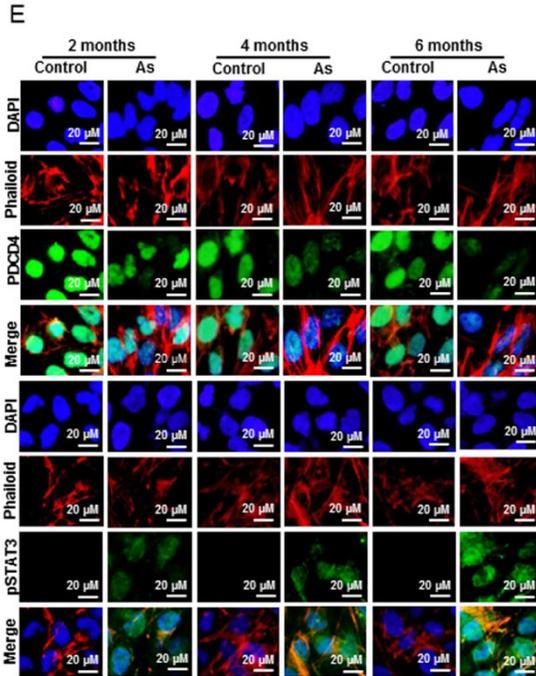


Figure 3. The arsenic-induced miR-21 increase and PDCD4 suppression contribute to malignant cell transformation. BEAS-2B cells were maintained in a medium containing various concentrations of arsenic (0.1, 0.25 and 0.5 μ M) for 6 months. (A,B) Cells were cultured in 0.35% soft agar for 5 weeks and number of colonies in the entire dish counted. (A) Representative images of control (left panel) and arsenic-treated (right panel) colonies. (B) Colony number increased in a dose-dependent manner. (C) The relative miR-21 level, determined by Taqman real-time PCR, increased in a time- and dose-dependent manner. (D) Total cell lysates were prepared for western blot analysis after 2, 4 and 6 months exposure to arsenic using specific antibodies against PDCD4, p47phox, pSTAT3 and STAT3. Apparent protein levels for PDCD4 decreased and P47phox and pSTAT3 increased in a time- and dose-dependent manner. (E) Representative images of fluorescence immunostaining for PDCD4 and pSTAT3 after 2, 4 and 6 months exposure to arsenic, and confirm results from western blot analysis. Data presented in the bar graphs are the mean \pm SD of three independent experiments. *Indicates a statistically significant difference compared to control with $p < 0.05$.

7.3c. *Data provided.* See Appendix 119.

Fig. 3E

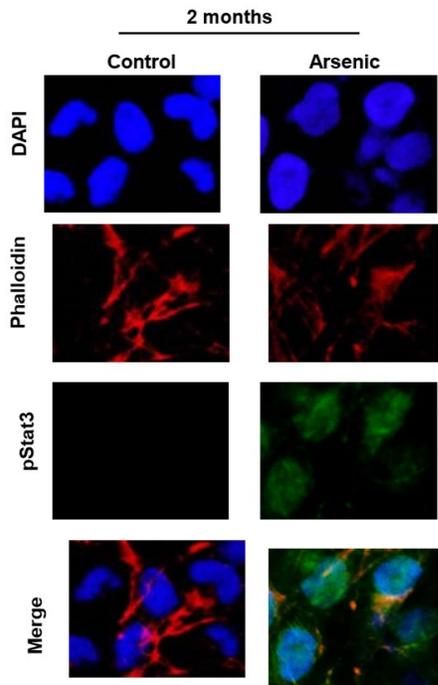


Fig. 3E

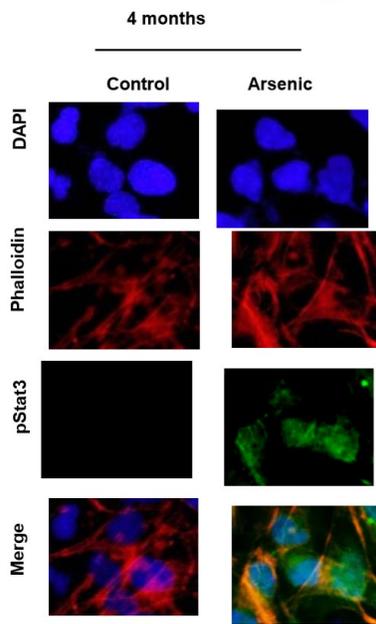
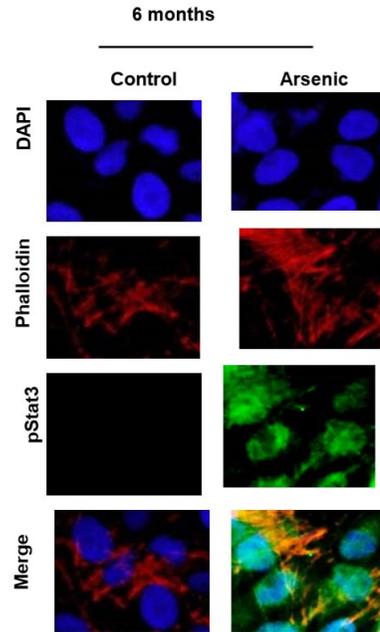
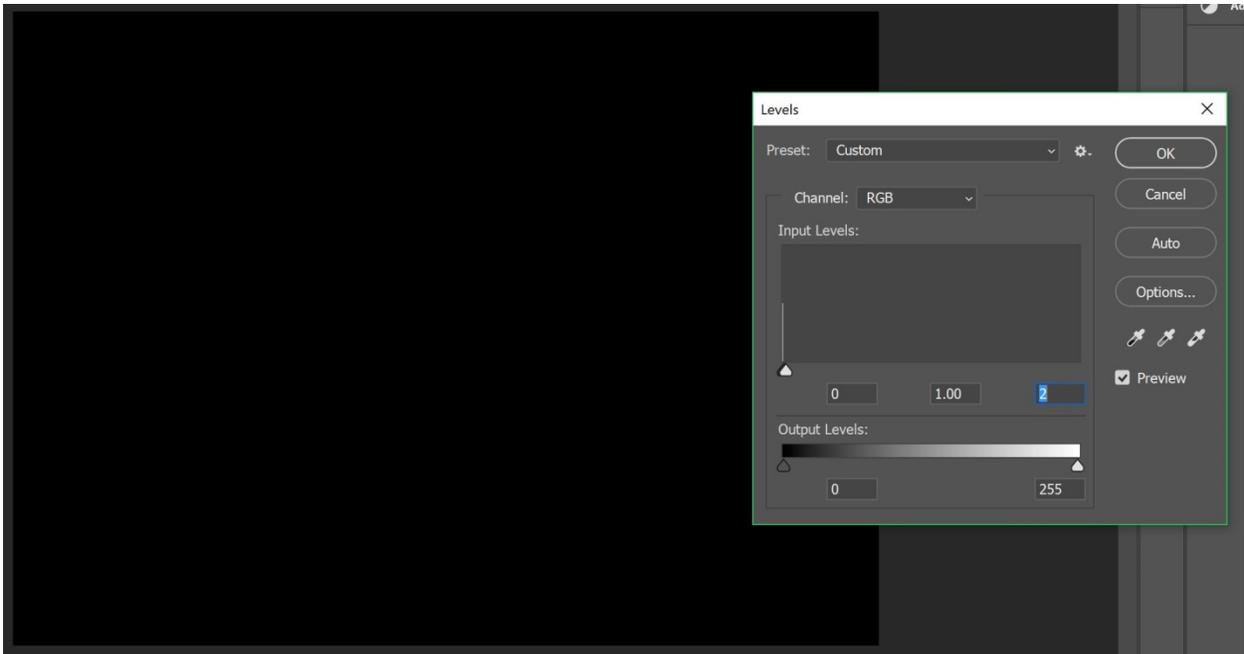


Fig. 3E

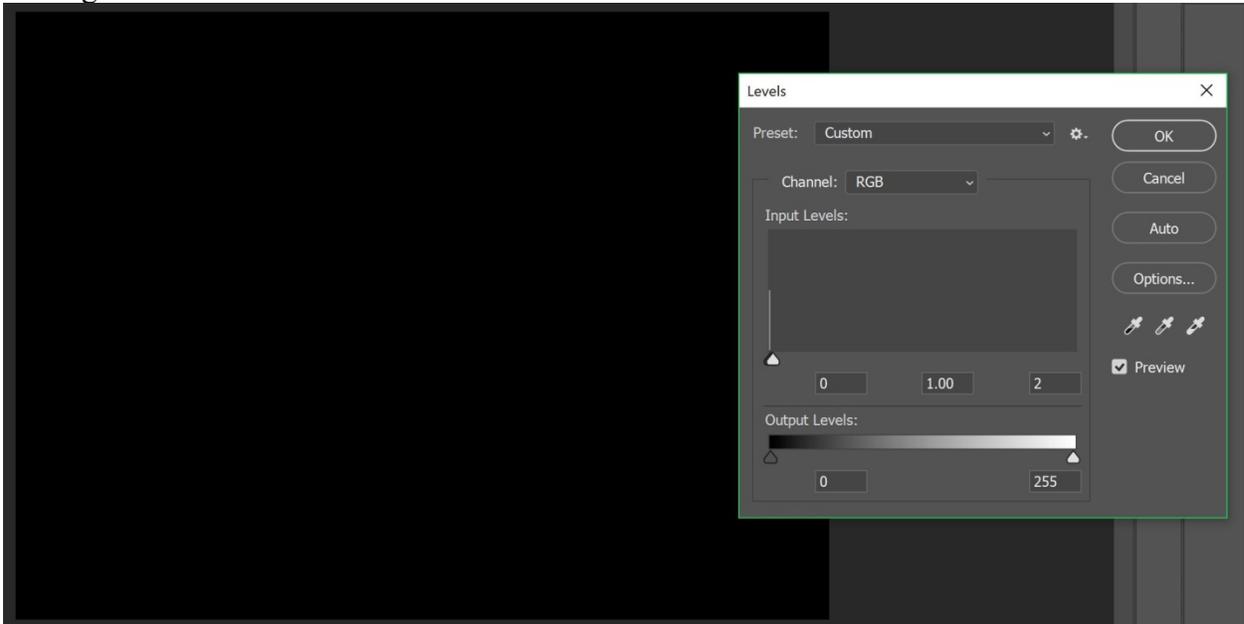


7.3d. *Our analysis.* In our analysis (and also as indicated by the Federal Office of Research Integrity), the panels for pSTAT3 2-month Control, 4-month Control, and 6-month Control in the published Fig. 3E did not have any pixel data in them. The three images used for the panels were black. A caveat to our analysis is that the committee had to work with a PDF of the original data rather than the original image files, as the Respondents did not provide the raw images. The Respondents included the following statement with the provided data: “Dr. Pratheeshkumar Poyil states that there was no basal expression of phosphor Sta3 found in Beas-2B cell by IF. Unstained cells cannot be detected even after increase brightness.” (Appendix 119)

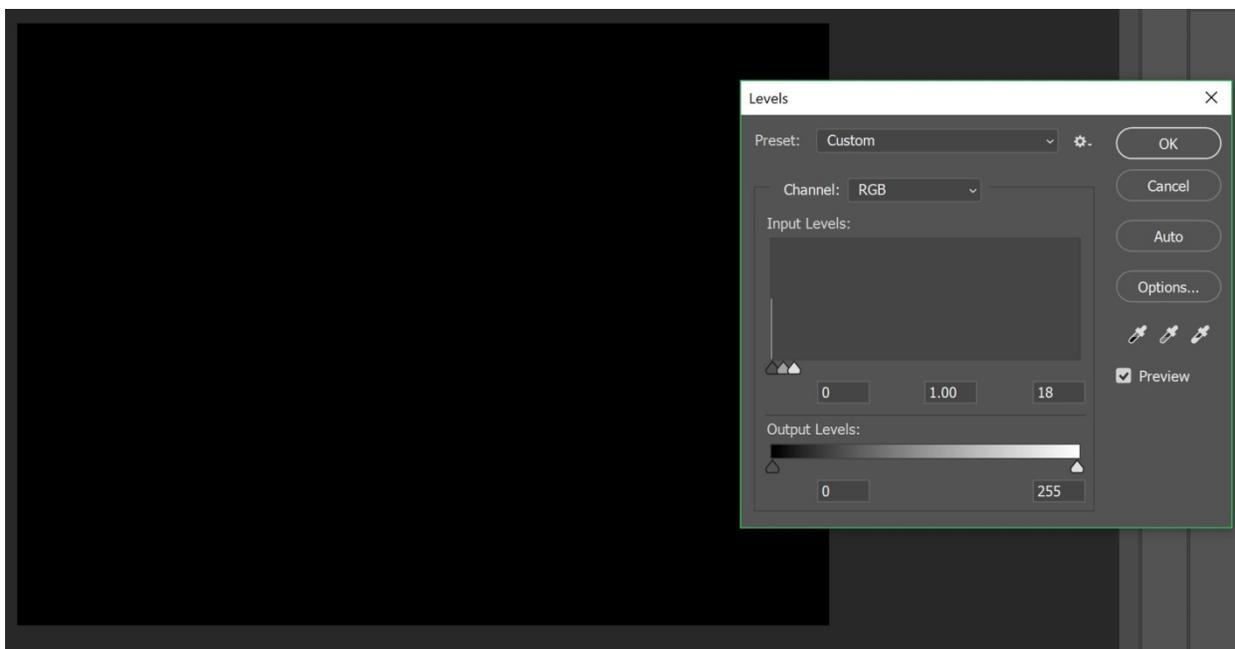
For Fig. 3E 2 months:



For Fig. 3E 4 months:



For Fig. 3E 6 months:



- 7.3e. *Relevant interview questions/comments.* The committee only asked questions about the scale bars in these figures and did not address the three black panels.
- 7.3f. *Conclusion for this specific example.* The three black panels in Fig. 3 did not show any evidence of fluorescence signal or background noise, and thus had no data in it. This is identical to the images of blank slides that Dr. Kim took during the committee's visit to the lab. Using a black box instead of actual data is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

8. Data fabrication and falsification:

Description: The committee defined this category as presentations of data that have been manipulated in such a way as to change the interpretation of the original data.

Overall summary: 2 figures from grants and 1 figure panel from manuscripts. (sections 8.1-8.3).

Overall summary of interviews and relevant comments: Since each incident was unique, there are no relevant general comments from the interviews. See specific comments in sections 8.1-8.3 below.

Overall conclusions: This category contains misrepresentation of data in grants and published figures. Specifically, the committee found manipulations of images that changed interpretations of the experiment and presented falsified view of the results. The motivations for these alterations are not uniformly clear. In one case, it appears that the data were falsified to support the effect of a drug. In another case, the falsified data actually weakened the arguments supporting the hypothesis. The committee was unable to decide whether this was a conscious effort to deceive or

just careless handling of the experimental data and figure construction. Regardless of the motivation, this represents scientific misconduct and was done intentionally.

During the investigation, one of the Respondents, Dr. Kim, provided falsified documents to support his narrative of how the experiment was performed and how Fig. 4C of M1 was generated (section 8.3).

Specific examples for “data fabrication”:

- 8.1. G7: Grant Zhang_3200001897 = 1R01CA228236-01A1 (01/01/2018-05/31-2023) (Appendix 012): Fig. 8, which is the same as Fig. 3D in M1 Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).
- 8.1a. *Date requested:* October 18, 2018 by the committee *via* email.
Date received: October 26, 2018 by Respondents *via* flash drive.
- 8.1b. *Original figure from Appendices 012, page 64 and 013, page 395.*

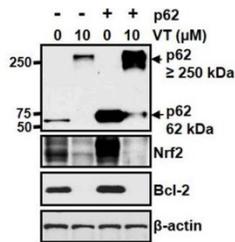


Fig 8 Verteporfin decreases Nrf2 through inhibition of p62. PC3 cells were transfected with pcDNA3.1/p62 plasmid for 48 h followed by verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments.

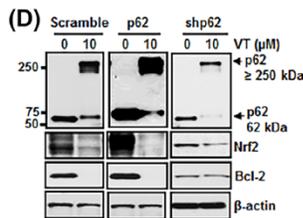
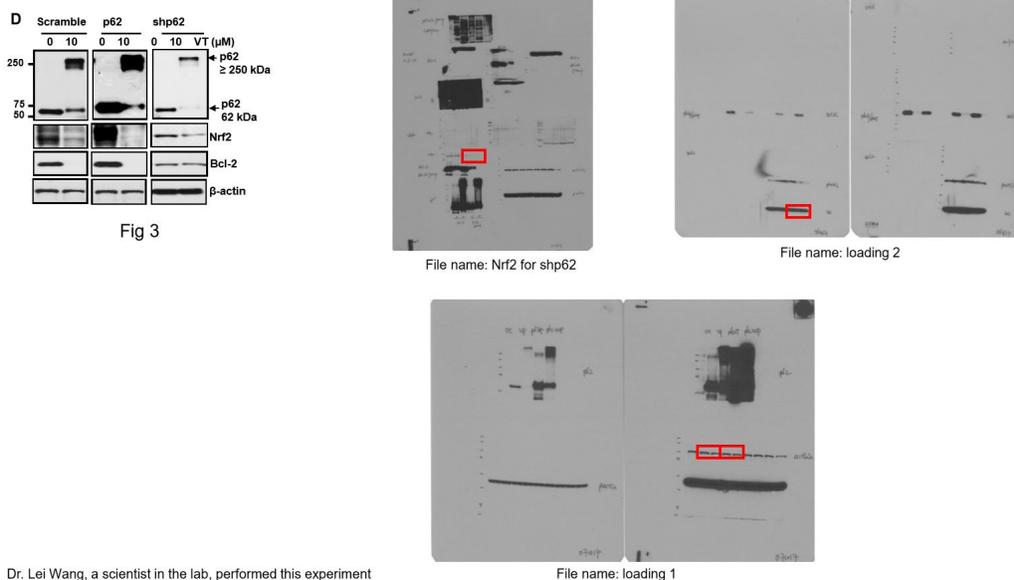
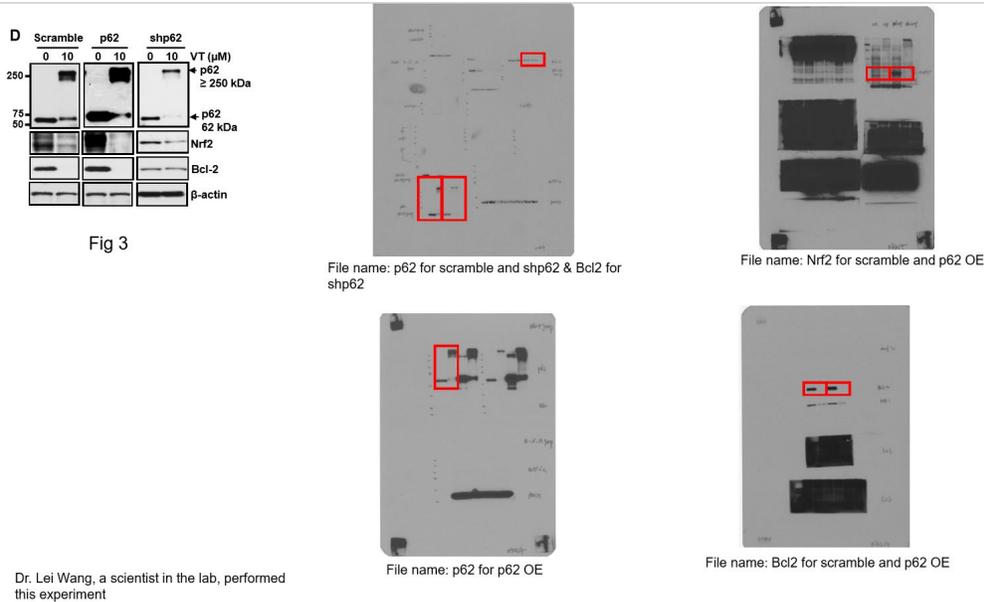
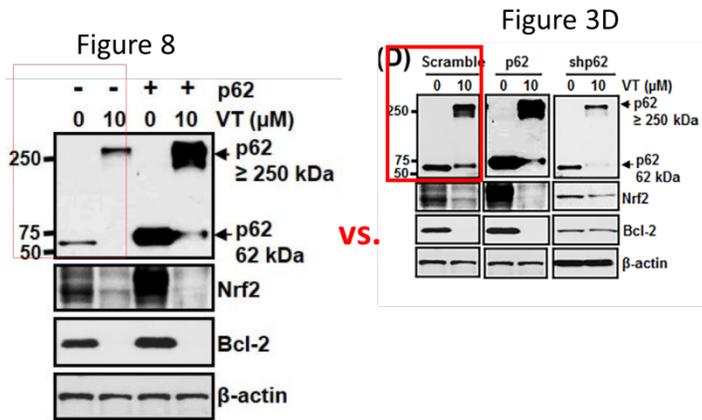


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μM hydroxychloroquine (HCQ), 10 μM Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μM Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μM and 10 μM Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μM 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean ± SE (n = 6). * and #, P < 0.05 compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μM of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

- 8.1c. *Data provided.* See Appendix 133.



- 8.1d. *Our analysis.* The committee examined Fig. 8 from grant G7 and Fig. 3D from M1 and concluded that they represented the same experimental data (same blots for Nrf2, Bcl2, and β-actin, which was actually α-actinin, see Section 2.6 above). The blots for p62 used for the “scrambled” lanes in Figure 3D were not the same as the p62- lanes in Figure 8. Only the last two of these p62 lanes were used to create Fig. 3D, and a set of two new lanes were added as lanes 1 and 2 (“scrambled”). New blots for Nrf2, Bcl2 and β-actin were not replaced in Fig. 3D, thus the data presented in Fig. 3D is not from the same set of samples (or even experiment, if indeed a scramble virus was used as a control in Fig. 3D).



8.1e. *Relevant interview questions/comments.* Not applicable.

8.1f. *Conclusion for this specific example.* This compilation of data is not consistent between the two figures and thus one or the other is a misrepresentation of how the experiments were performed. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification and fabrication.

8.2. G7: Grant Zhang_3200001897 = 1R01CA228236-01A1 (01/01/2018-05/31-2023) (Appendix 012): Fig. 12, which is the same as Fig. 3G in M1 Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013).

8.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

8.2b. *Original figure from Appendices 012, page 65 and 013, page 395.*

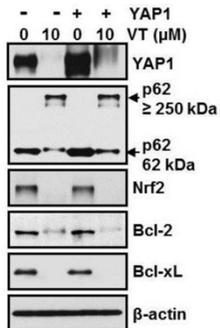


Fig 12 Verteporfin inhibits p62 and its downstream proteins independent of YAP1. PC3 cells were transiently transfected with pCMV6-GFP/YAP1 followed by treatment with or without verteporfin. At 24 h, whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments.

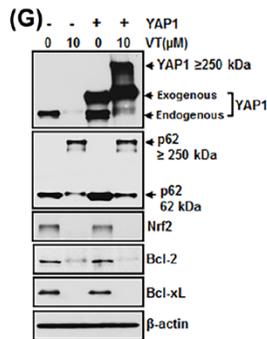
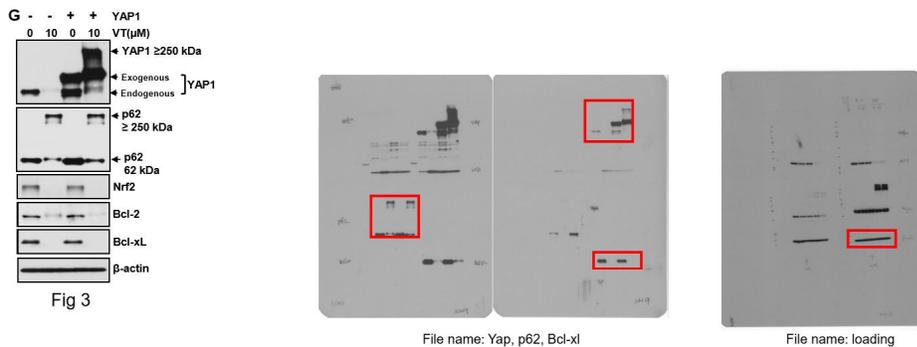


FIGURE 3 Verteporfin decreases autophagy and pathways downstream of p62 including activation of constitutive Nrf2, its target proteins, and ROS level. (A) Verteporfin inhibited basal levels of LC3-II and those enhanced by HCQ and bafilomycin A1 in PC-3 and LNCaP cells. PC-3 or LNCaP cells were treated with 10 μ M hydroxychloroquine (HCQ), 10 μ M Verteporfin (VT), and 10 nM bafilomycin A1 (Baf A1) alone or in combination for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B) Verteporfin decreased LC3 puncta formation. PC-3 or LNCaP cells were starved overnight and then treated with 10 μ M Verteporfin for 24 h. The LC3 puncta were visualized using fluorescence microscopy. Photomicrographs demonstrate immunofluorescence staining for LC3 puncta formation. (C) Verteporfin inhibited p62 downstream signaling. PC-3 cells were treated with 5 μ M and 10 μ M Verteporfin (VT) for 24 h. Whole-cell lysates were collected for immunoblotting analysis. (D) Verteporfin decreased Nrf2 through inhibition of p62. PC-3 cells were either transfected with pcDNA3.1/p62 or p62 shRNA plasmid for 48 h followed by Verteporfin (VT) treatment for 24 h. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (E) Verteporfin increased ROS generation in PC-3 cells. PC-3 cells were transfected with p62 shRNA or scramble for 24 h followed by treatment with various doses of Verteporfin for 6 h followed by staining with 10 μ M 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate ethyl ester (DCFDA) for 30 min. Fluorescence intensity was measured by flow cytometry. The results are shown as mean \pm SE (n = 6). * and #, $P < 0.05$ compared to control without treatment in scramble cells and p62 shRNA transfected cells, respectively. (F) Verteporfin induced apoptosis in PC-3 cells. PC-3 cells were treated with 10 μ M of Verteporfin for 24 h. Apoptosis was measured by Annexin V-FITC/PI assay. (G) Inhibition of p62/Nrf2 signaling by Verteporfin independent of YAP1. PC-3 cells with transient transfection with YAP1 overexpressing plasmid were treated with Verteporfin for 24 h. Whole-cell lysates were harvested for immunoblotting analysis. The results represent three independent experiments

8.2c. *Data provided.* See Appendix 112.



8.2d. *Our analysis.* It should be noted that these figures were also discussed in section 3.3. Fig. 12 in the grant G7 and Fig. 3G in manuscript M1 show data from the same experiment. To create Fig. 12 in G7, the YAP1 portion of the image in Fig. 3G from M1 was cropped to exclude two higher molecular weight bands corresponding to the exogenously expressed YAP1 (GFP-YAP1) and some undefined, high-molecular-weight form of YAP1. Based on the data in Fig. 3G, the drug verteporfin reduces endogenous YAP1 levels, yet has no effect on the levels of exogenously expressed GFP-YAP1. The image manipulation hides this experimental fact in the grant Fig. 12.

8.2e. *Relevant interview questions/comments.* See section 3.3.

8.2f. *Conclusion for this specific example.* The grant figure misrepresents the effect of the drug on the exogenously-expressed GFP-YAP1. In Fig. 12, it appears that verteporfin affects both endogenous and exogenously-expressed YAP1 because only one band for YAP1 was shown. However, Fig. 3 demonstrates that the levels of exogenously-expressed protein were resistant to verteporfin. It is unclear why the Respondents chose to present the data this way in Fig. 12 since showing all YAP1 bands would have supported their hypothesis about p62's independence on YAP1; the caveat being that the exogenous GFP-YAP1 and the endogenous YAP1 proteins function similarly in the cells. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification and fabrication.

8.3. M1: Wang L, Kim D, Wise JTF, Shi X, Zhang Z, and DiPaola RS. (2018). p62 as a therapeutic target for inhibition of autophagy in prostate cancer. *The Prostate*, 78, 390-400 (Appendix 013): Fig. 4C

8.3a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

8.3b. *Original figure from Appendix 013, page 396.*

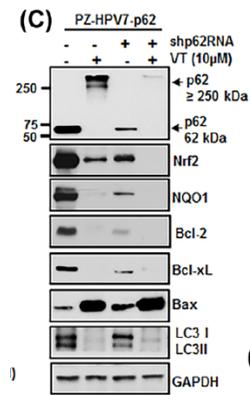
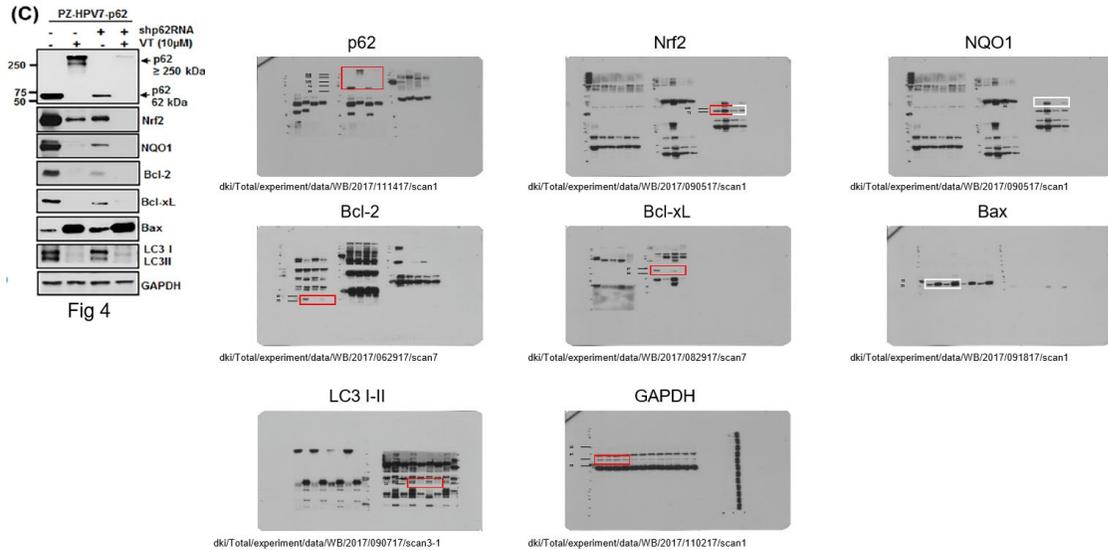


FIGURE 4 Overexpression of p62 in normal prostate epithelial cells led to apoptosis resistance and increased autophagy flux, and Verteporfin exhibited inhibitory effects. (A) Overexpression of p62 in normal prostate epithelial cells led to a profile consistent with apoptosis resistance. Normal prostate epithelial RWPE-1 and PZ-HPV7 cells were transfected with pcDNA3.1/p62 plasmid followed by antibiotic selection for two months. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B-D) Verteporfin caused p62 crosslink and decreased Nrf2 and autophagy. (B) p62-expressing RWPE-1 and PZHPV cells were treated with Verteporfin for 24 h. (C) p62-expressing PZ-HPV cells were transiently transfected with p62 shRNA for 24 h followed by treatment with Verteporfin. (B) and (C) Whole-cell lysates were harvested for immunoblotting analysis. (D) PZ-HPV7 cells with or without overexpression of p62 or p62 shRNA were transfected with GFP-mCherry-LC3 plasmid followed by treatment with Verteporfin (10 µM) for 24 h. The cells were visualized using fluorescence microscopy. The results are representative of three independent experiments. (E) and (F) Verteporfin increased ROS generation and apoptosis. p62-expressing PZ-HPV7 cells were transfected with p62 shRNA for 24 h followed by Verteporfin treatment. After 24 h, the cells were collected for ROS measurement (E) and apoptosis analysis (F)

8.3c. *Data provided.* See Appendix 134.

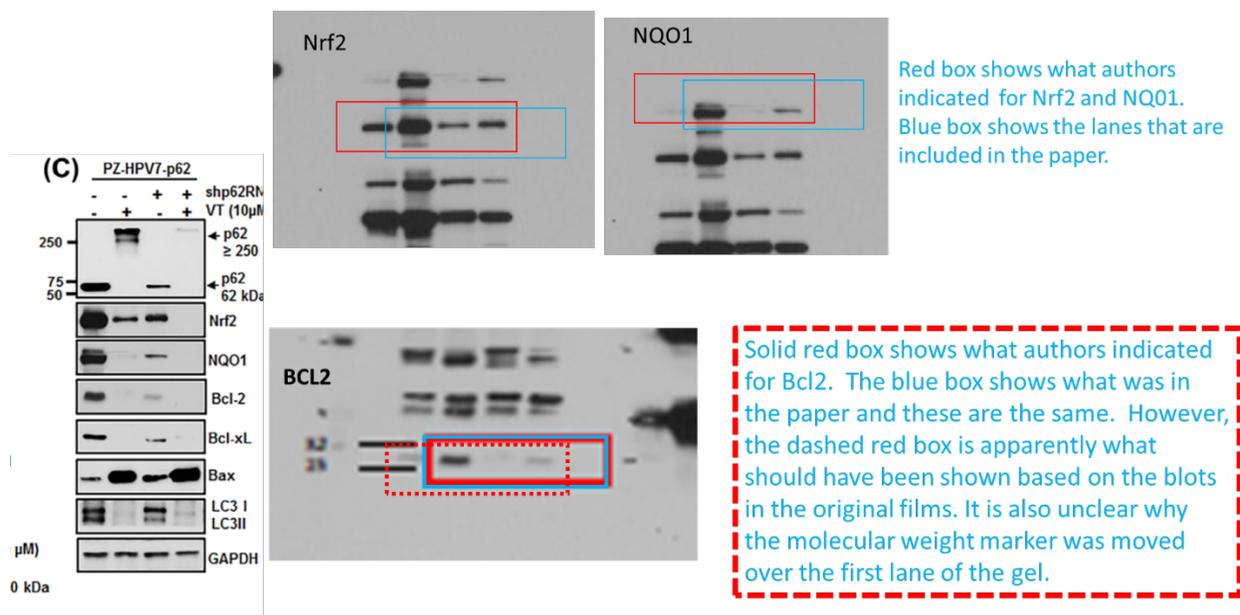


8.3d. *Our analysis.* The data provided indicate that the gels for this experiment were loaded in sets of four samples. For the Nrf2 and NQO1 western blots, four lanes containing bands were marked in the original scans by the Respondents (red boxes in the figure). However, the lanes in the published figure for these two western blots do not match what was indicated by the Respondents. It appeared that the published figure was shifted to the right by one lane, so that the first lane was excluded and a blank lane was added as the fourth (blue box in the figure). Based on the appearance of the images of the original western blots and the fact that all samples for this experiment were loaded as sets of four, the committee is not convinced that the included blank lane contained any sample. The committee noted that background bands were presented in the first four lanes of the western blot but did not appear in the fifth lane that the Respondents claimed contained a sample.

For the western blots for Bcl-2, the Respondents indicated four lanes, including a blank lane as the fourth (solid red box in the figure). What the Respondents indicated matches what was published; however, what was published does not match the four lanes that appear to contain sample in the provided image (dotted red box in the figure). As with the other two blots, this included fourth lane did not appear to contain a sample. The committee noted that background bands were presented in the first four lanes of the western blot but did not appear in the fifth lane that the Respondents claimed contained a sample.

Since all Respondents indicated that Dr. Kim generated the data and the figure (below), he was specifically asked about how the samples for this experiment were analyzed and then the data presented. After repeated questioning, Dr. Kim maintained that there were five samples loaded and the first lane represented a standard control that he chose not to include in the published figure (Appendix 063 (Kim interview II, pages 19-28)). After the interview, Dr. Kim submitted more information to support his contention (Appendix 069). This material included screenshots of two slides (slides A and B) from a PowerPoint file (labeled-“blot cutter 1”) that showed how the samples were loaded on the gels for this experiment (Appendix 135; Slide A is page 3 of that Appendix, and Slide B is page 4 of that Appendix). Also included was the file path to the specific PowerPoint file. Since the

committee had sequestered the hard-drives where this file was stored, it examined the file and could not find the two slides in the PowerPoint file indicated by the file path. Based on the sequence of the slides in the sequestered file, it appears that the two slides (slides A and B) had been inserted at a later date. A detailed analysis of these findings is presented in Appendix 136.



8.3e. *Relevant interview questions/comments.* Drs. Shi, Wang and Kim stated that Dr. Kim performed the experiments for and prepared Fig. 4 (Appendices 052 (Shi interview, pages 36-37); 058 (Wang interview, page 26); 063 (Kim interview II, page 5), and 054 (Zhang interview, page 31)). He gave the final figure in PowerPoint to Dr. Zhang (Appendix 063 (Kim interview II, page 7)). Dr. Wang only saw the data as a PowerPoint (Appendix 058 (Wang interview, page 56)). Dr. DiPaola agreed that the figures do not match the original data (Appendix 060 (DiPaola interview II, pages 19-20)). He further agreed that if there were only four lanes it would change the conclusions of Figure 4 (Appendix 060 (DiPaola interview II, pages 21-22)). Dr. Shi stated that the data for Nrf2 and NQO1 do not appear to match (Appendix 052 (Shi interview, pages 83-84)). Dr. Zhang was unclear about what was loaded in the fifth lane (Appendix 054 (Zhang interview, pages 73-75)). Dr. Zhang stated that the red boxes in the data provided do not make sense. She thought that Kim mis-boxed the lanes (Appendix 054 (Zhang interview, pages 76-77)). Dr. Kim's specific comments on Fig. 4C are noted in the analysis section above.

8.3f. *Conclusion for this specific example.* The committee concluded that the published figure does not match the original data. The shift of the lanes for the Nrf2, NQO1 and Bcl-2 western blots changes the interpretation of the experiment. Based on the original scans, verteporfin increases expression of Nrf2, NQO1 and Bcl-2. However, as presented in the published figure, the drug decreases the level of all three proteins. Thus, the data presented in the publication are diametrically opposed to the results of the experiment. The committee would have not uncovered this discrepancy had it not had access to the original data. Additionally, based on the information recovered, documents related to this experiment

FIGURE 4 Overexpression of p62 in normal prostate epithelial cells led to apoptosis resistance and increased autophagy flux, and Verteporfin exhibited inhibitory effects. (A) Overexpression of p62 in normal prostate epithelial cells led to a profile consistent with apoptosis resistance. Normal prostate epithelial RWPE-1 and PZ-HPV7 cells were transfected with pcDNA3.1/p62 plasmid followed by antibiotic selection for two months. Whole-cell lysates were collected for immunoblotting analysis. The results represent three independent experiments. (B-D) Verteporfin caused p62 crosslink and decreased Nrf2 and autophagy. (B) p62-expressing RWPE-1 and PZHPV cells were treated with Verteporfin for 24 h. (C) p62-expressing PZ-HPV cells were transiently transfected with p62 shRNA for 24 h followed by treatment with Verteporfin. (B) and (C) Whole-cell lysates were harvested for immunoblotting analysis. (D) PZ-HPV7 cells with or without overexpression of p62 or p62 shRNA were transfected with GFP-mCherry-LC3 plasmid followed by treatment with Verteporfin (10 μ M) for 24 h. The cells were visualized using fluorescence microscopy. The results are representative of three independent experiments. (E) and (F) Verteporfin increased ROS generation and apoptosis. p62-expressing PZ-HPV7 cells were transfected with p62 shRNA for 24 h followed by Verteporfin treatment. After 24 h, the cells were collected for ROS measurement (E) and apoptosis analysis (F)

9.1c. *Data provided.* See Appendices 069 and 135.

9.1d. *Our analysis.* After the interview, Dr. Kim submitted more information *via* email on February 21, 2019 to support his contention (Appendices 069 and 132). This material included screenshots of two slides (slides A and B) from a PowerPoint file (labeled-“blot cutter 1”) that showed how the samples were loaded on the gels for this experiment (Appendix 135; Slide A is page 3 of that Appendix, and Slide B is page 4 of that Appendix). Also included was the file path to the specific PowerPoint file. Since the committee had sequestered the hard-drives where this file was stored, it examined the file and could not find the two slides in the PowerPoint file indicated by the file path. Based on the sequence of the slides in the sequestered file, it appears that the two slides (slides A and B) had been inserted at a later date. A detailed analysis of these findings is presented in Appendix 136.

9.1e. *Relevant interview questions/comments.* See Appendix 069.

9.1f. *Conclusion for this specific example.* Based on the information recovered, Dr. Kim provided falsified documents to support his narrative of how the experiment was performed and how Fig. 4C of M1 was generated. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data falsification.

9.2. G3: Grant Shi_3200001792 = 1R01ES029378-01 (04/01/2018-03/31/2023) (Appendix 008):
Fig. 6

9.2a. *Date requested:* October 18, 2018 by the committee *via* email.

Date received: October 26, 2018 by Respondents *via* flash drive.

9.2b. *Original figure from Appendix 008, page 53.*

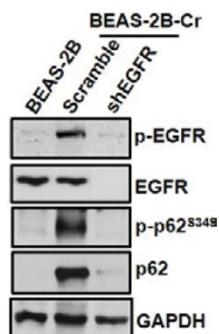


Fig. 6 EGFR upregulates p62 in Cr(VI)-transformed cells. Passage-matched normal cells (BEAS-2B) and Cr(VI)-transformed cells (BEAS-2B-Cr) with (shEGFR) or without (Scramble) EGFR knockdown were cultured in 10-cm culture dishes. The cells were harvested and whole protein was isolated for immunoblotting analysis. The results are representative of three independent experiments.

- 9.2c. *Data provided.* See Appendix 065 and 137.
- 9.2d. *Our analysis.* As a result of the interview, where Dr. Shi was specifically questioned about this figure, he provided data (to Dr. Crocker 2/12/19 *via* email, Appendix 065) that he claimed showed that the blank lane used in the EGFR panel of the figure corresponded to a lane where sample had been loaded (section 3.1 for description of the figure irregularity). The data provided was a gel of the actin loading control that contained 10 lanes. It should be noted that in the grant figure, GAPDH and not actin was used as loading control. Dr. Shi provided the file path to the images of the scan of the actin loading control. Since the committee had sequestered the hard-drives where this file should have been stored, it examined the drives and could not find the file indicated by the file path.
- 9.2e. *Relevant interview questions/comments.* Dr. Zhang stated that Dr. Kim provided the PowerPoint for this figure and labeled it (Appendix 054 (Zhang interview, page 70)). This was further confirmed in the material provided after the interview by Dr. Shi (Appendix 137)
- 9.2f. *Conclusion for this specific example.* Based on the information recovered and the fact that the file could not be found on the sequestered hard-drive, the committee concludes that the Respondents provided them with fabricated documents (Appendix 134) to support their narrative of how the experiment was performed and how Fig. 6 of G3 was generated. This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

10. Retracted manuscripts: 3 manuscripts. *As requested by the Federal Office of Research Integrity, in this category analysis of all figures in retracted manuscripts is performed.*

Overall summary of issues identified in all three retracted JBC manuscripts: For the analysis of the three retracted JBC manuscripts, the committee first scrutinized the problems identified in the JBC retraction notices (Appendices 055, 056, and 057). After request from the committee for raw data for all figures in the three retracted manuscripts on April 16, 2019 *via* email (Appendix 078), the Respondents provided the data in electronic format and/or lab notebook.

None of the electronic data provided by the Respondents were true raw data as only JPG images of the following were provided, (1) cropped western blot bands that were exactly what was in the published figures, (2) flow-cytometry plots, (3) ESR spectra, and (4) fluorescence images with no file paths. None of these files had relevant metadata associated with them. Additionally, Excel files were provided for some bar graph figures, but they did not have the creation date or the date the experiment was performed on, and there was no file path indicated. Thus, the committee cannot confirm the validity of the electronic data.

For the data provided in lab notebook format (8 lab notebooks), the Respondents indicated pages potentially related to published figures by post-it notes. In few cases the post-it notes pointed to an exact figure in one of the three retracted manuscripts, while in most cases the post-it notes contained only information on the manuscript to which the data could potentially be related. Since the data in the lab notebook and the post-it notes placed by the Respondents did not provide information on the experiments done, the committee could not validate these data despite analyzing them with due diligence. For the cases where a specific figure/manuscript were indicated, the committee investigated whether the data provided were the ones published (analysis below; in some cases, the data provided matched the publication, in others they did not).

In some instances, the committee could not match the raw data provided to a published figure in these manuscripts. The ESR spectra in lab notebook indicated by the Respondents did not match the ones published in the three manuscripts (the noise in the raw data is significantly smaller compared to the one in the published figures). The Respondents marked a large number of flow-cytometry scans (data with values coming from the instrument) as belonging to the three manuscripts, without a clear indication of which exact scans were used. As the metals indicated in the scans did not always match those in the manuscripts, the committee could not verify the validity of the data. The Respondents also provided lab notebook pages with prepared figures (without raw data for the panels) that match the flow-cytometry images in the three JBC manuscripts.

When looking at PubPeer, a data integrity watchdog site, there are manuscripts with noted irregularities going back to 2004. While beyond our investigative horizon (~6 years), it is apparent that some of the image manipulations and load control duplications have been occurring in manuscripts from this group for quite some time. This past pattern matches much of what the committee has uncovered in more recent publications.

Additional retracted publications: The committee did a PubMed search for “Shi X AND retracted” and found 7 hits. In addition to the three manuscripts that Federal Office of Research Integrity asked the committee to investigate, the committee found one additional manuscript with Dr. Shi as the corresponding author that was retracted in 2002 because of overlap with another manuscript published by the same group (Chen et al., PNAS, 2002 below). It is important to note that the committee asked Dr. Shi during the interview on 01/17/2019 if he had any manuscripts retracted other than the three JBC manuscripts that were retracted on September 2018: question = “Have you ever had to retract a paper before?” and he answered “No, never” (Appendix 052 (Shi interview, page 94)).

[J Biol Chem](#). 2002 Jul 3. [Epub ahead of print]

RETRACTED: Ubiquitination and degradation of Cdc25C contribute to As(III)-induced G2/M cell cycle arrest.

[Chen F¹](#), [He ML](#), [Bower J](#), [Sbarra D](#), [C M Lin M](#), [Kung HE](#), [Shi X](#).

[Author information](#)

Abstract

The first author has retracted this paper because of an overlap with Chen, et al, (2002), PNAS 99: 1990-1995, DOI 10.1073/pnas.032428899.

PMID: 11880363 DOI: [10.1074/jbc.M107813200](#)

[Free full text](#)



[Proc Natl Acad Sci U S A](#). 2002 Feb 19;99(4):1990-5. Epub 2002 Feb 12.

Arsenite-induced Cdc25C degradation is through the KEN-box and ubiquitin-proteasome pathway.

[Chen F¹](#), [Zhang Z](#), [Bower J](#), [Lu Y](#), [Leonard SS](#), [Ding M](#), [Castranova V](#), [Pivnicka-Worms H](#), [Shi X](#).

[Author information](#)

Abstract

Arsenite is a known human carcinogen that induces tumorigenesis through either a genotoxic or an epigenetic mechanism. In this study, the effect of arsenite on cell cycle regulation and the mechanisms that contribute to this effect were investigated. Treatment of the cells with arsenite suppressed cell proliferation and reduced cell viability in a dose- or time-dependent manner. Analysis of cell cycle profile and cell cycle regulatory proteins indicated that arsenite arrested the cell cycle at G(2)/M phase, partially through induction of cell division cycle 25 (Cdc25) isoform C (Cdc25C) degradation via ubiquitin-proteasome pathways. Mutation of the putative KEN box within the region 151 to 157 of human Cdc25C or treatment of the cells with a peptide competitor encompassing the KEN box partially inhibited arsenite-induced ubiquitination of Cdc25C. Thus, these results indicate that the regulated ubiquitination of Cdc25C may be involved in the arsenite-induced proteolytic down-regulation of Cdc25C activity in the G(2)/M phase of the cell cycle and suggest a link between cell cycle and the carcinogenic effects of arsenite.

PMID: 11842186 PMCID: [PMC122307](#) DOI: [10.1073/pnas.032428899](#)

[Indexed for MEDLINE] [Free PMC Article](#)

Detailed analysis by the committee of issues identified in all three retracted JBC manuscripts:

- 10.1. R1: Son, Y.O., Pratheeshkumar, P., Divya, S.P., Zhuo Zhang, Z., and Shi, X. Nuclear factor erythroid 2-related factor 2 enhances carcinogenesis by suppressing apoptosis and promoting autophagy in nickel-transformed cells. *Journal of Biological Chemistry*, 292, 8315-8330, 2017 (Appendix 055): All figures.

This manuscript was retracted on September 7, 2018.

- 10.1a. *JBC retraction notice*: The following three issues were identified in the JBC retraction notice (Appendix 055):

- (i) The image from NiT cells treated with Ni²⁺ was previously published in Figure 5H of Son *et al.* 2015, representing different experimental conditions.
- (ii) The first two lanes of the GAPDH immunoblot from Figure 3H were reused in Figures 4A and 7A.

(iii) The ki67 and H&E panels from normal tissue in Figure 8I were reused from Figure 7C of Son et al. 2015 without attribution.

10.1b. *The committee's analysis of issues identified in the retraction notice from JBC:*

Issue (i): The image from NiT cells treated with Ni²⁺ was previously published in Figure 5H of Son et al. 2015, representing different experimental conditions.

Fig. 2F from retracted manuscript R1:

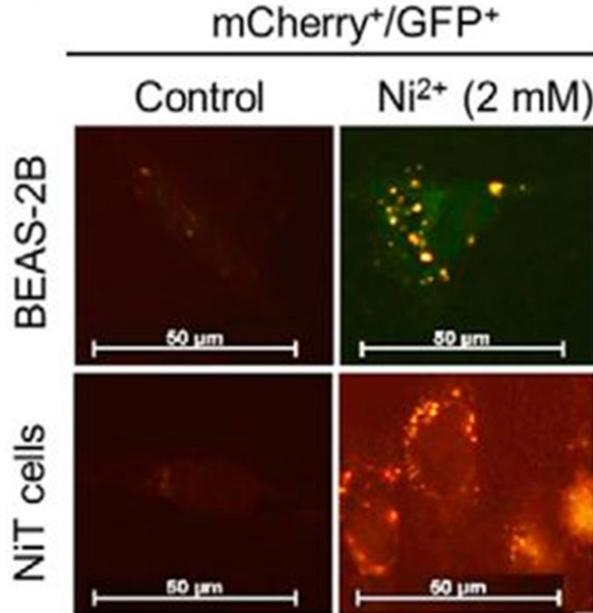
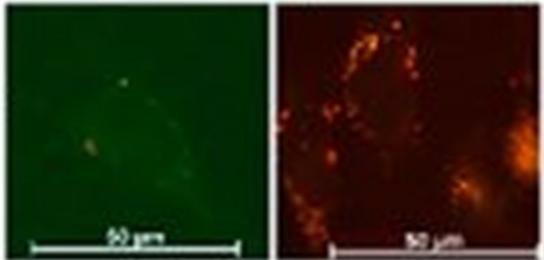


Fig. 5H from Son et al., JBC, 2015 (retracted manuscript R2):



The committee confirmed that the bottom right panel in Fig. 2F is the same image used in Fig. 5H of Son et al., 2015, representing different experimental conditions. In Fig. 2F, the panel represents NiT cells treated with 2 mM Ni, whereas in Fig. 5H of Son et al., 2015, the panel is claimed to represent cells treated with 20 μM As. This is an example of intentional data falsification.

Issue (ii): The first two lanes of the GAPDH immunoblot from Figure 3H were reused in Figures 4A and 7A.

Fig. 3H from retracted manuscript R1:

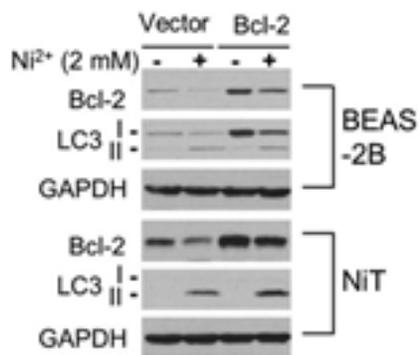
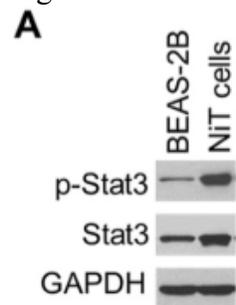


Fig. 4A from retracted manuscript R1:



Fig. 7A from retracted manuscript R1:



The committee confirmed that the first 2 lanes in the GAPDH gel of Fig 3H (Son, 2017, Appendix 055) were reused in Figs. 4A and 7A. The three experiments represented were different. This is an example of intentional data falsification.

Issue (iii): The ki67 and H&E panels from normal tissue in Figure 8I were reused from Figure 7C of Son et al. 2015 without attribution.

Fig. 8I from retracted manuscript R1:

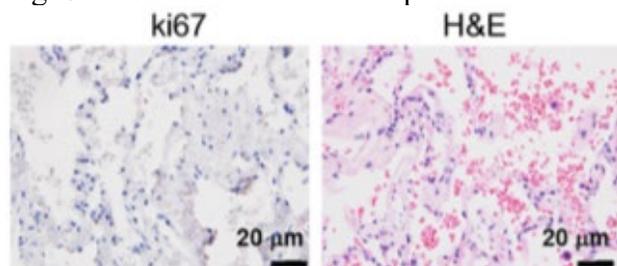
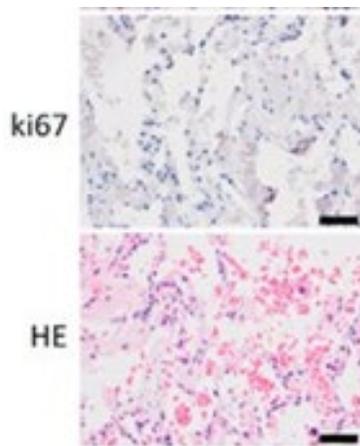


Fig. 7C from Son et al., JBC, 2015 (retracted manuscript R2, Appendix 056):



The committee confirmed that normal tissue images for the ki67 and HE staining in Figs. 8I (Son, 2017, Appendix 055) and 7C (Son, 2015, Appendix 056) are the same. This is an example of intentional data falsification.

Overall conclusions for these three issues: These issues represent a significant departure from accepted practices of the research community that were committed intentionally and these allegations are proven by a preponderance of evidence. The committee concluded that these are cases of data falsification.

10.1c. *Additional issues with the manuscript identified by the committee that the JBC retraction notice did not indicate:*

10.1c1. *Figures for which no electronic data were provided:* Figs. 1G (a duplicate was provided), 1H, 2E, 2G, lower right panel in 2F (discussed above as being duplicated from another manuscript), 3B, 3D, 3E, 3F, 3G, 3H, 4B, 4C, 4D, 4E, 4F, western blots for 5D and 5E, 5F, 5H, Excel data for 6A, 6B, numerous panels from 6E and 6F, 6H, 6K, 6L, 6M, 6N, western blots for 7B and 7C, 7G, 7J, images for 8D and 8H.

Not preserving raw data is a significant departure from accepted practices of the research community and is a breach of NIH and UK data retention policies (Appendices 091 and 092 that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

10.1c2. *Issues identified with electronic data provided:*

The scale bars in the published figure have been incorrectly represented. Based on the committee's examination of the magnifications of the images provided by the Respondents, the published image is not the same magnification as was the original. Thus the scale bars are incorrect. See a representative example below.

Fig. 8I, H&E normal as published:

H&E

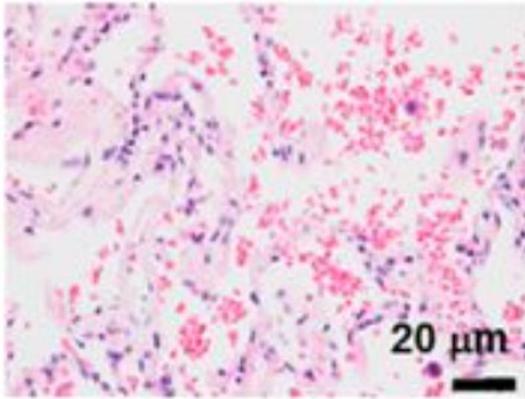
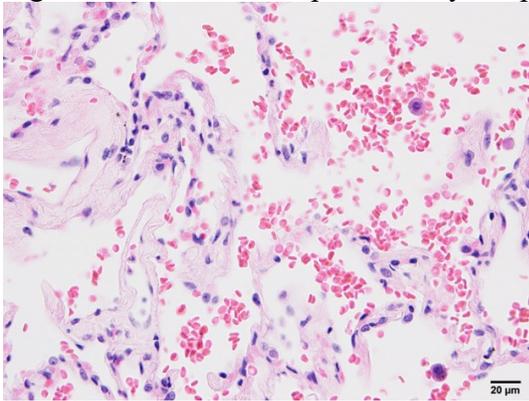


Fig. 8I, H&E normal as provided by Respondents:



This is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

10.1c3. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could verify:* Raw data in lab notebook format were provided for Figs. 1B, 1C, 8C, 6G, 6J, and the committee could confirm their validity.

10.1c4. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could **not** verify because of lack of labeling or because the data did not match the published figure:*

Fig. 1F: The committee could match the four images provided to the published figure. However, the percentage indicated in the published top right panel as 52.91 is actually 42.91 in the raw data in the notebook.

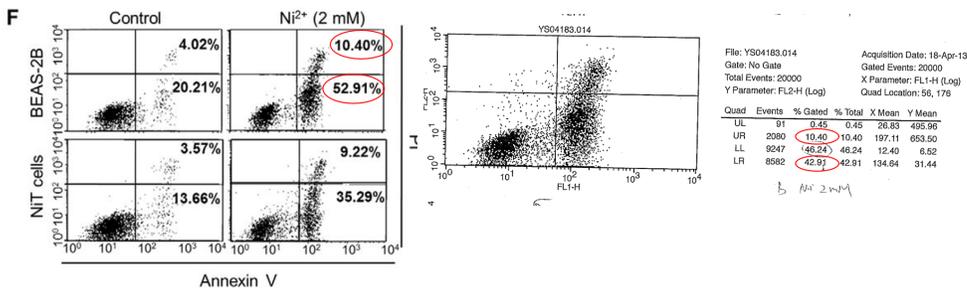
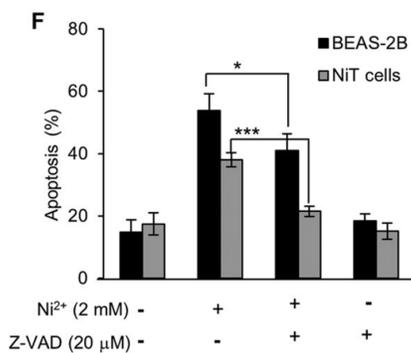


Fig. 3F: data provided in lab notebook do not match the published figure and the Ni treatment data is missing.



NiT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	CdT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA
0.55	7.2	2.61	7.74	1.23	0.77	1.19	8.98	2	8.52
11.48	27.32	19.5	29.64	8.09	12.39	13.01	37.9	14.62	39.47
12.44	34.52	22.11	34.38	9.32	13.16	14.2	46.66	16.62	47.99

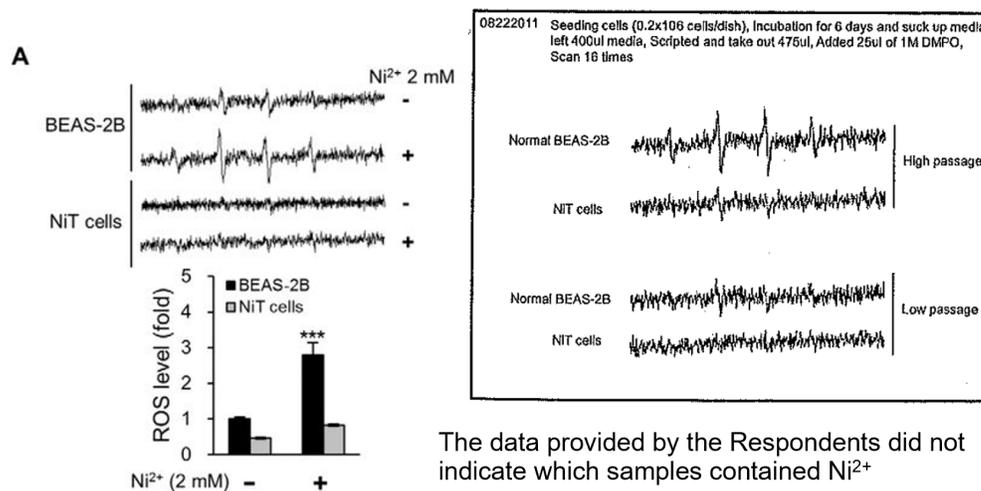
NiT-X1 Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	CdT-X1 Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA
1.61	8.3	3.84	8.68	1.88	0.99	0.45	2.03	1.57	3.37
12.44	34.95	10.59	32.11	7.39	15.41	7.67	23.09	5.55	18.81
14.05	43.25	14.43	40.79	9.27	16.39	8.12	25.72	7.12	22.18

NiT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	CdT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA
12.44	34.52	22.11	34.38	9.32	13.16	14.2	46.66	16.62	47.99
14.05	43.25	14.43	40.79	9.27	16.39	8.12	25.72	7.12	22.18
13.245	38.865	18.27	37.585	9.295	14.775	11.16	36.29	11.67	35.085

NiT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	CdT Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA
0.94	2.81	1.67	2.55	0.70	0.99	1.27	4.20	1.49	4.30
1.95	3.27	1.09	3.08	0.70	1.24	0.73	2.30	0.64	1.99
1.00	2.94	1.38	2.84	0.70	1.12	1.00	3.25	1.06	3.14

The Respondents indicate that the provided data table was incorporated into three different Figures from two distinct papers.

Fig. 6A: data provided in lab notebook do not match the published figure.



The data provided by the Respondents did not indicate which samples contained Ni²⁺

Since the raw data provided did not match the figures published, not preserving raw data is a significant departure from accepted practices of the research community and is a breach of NIH and UK data retention policies (Appendices 091 and 092) that was committed intentionally and

this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

- 10.2. R2: Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Divya, S.P., Xu, M., Luo, J., Chen, G., Zhang, Z. and Shi, X. Antioncogenic and oncogenic properties of Nrf2 in arsenic-induced carcinogenesis. *Journal of Biological Chemistry*, 290, 27090-27100, 2015 (Appendix 056): All figures.

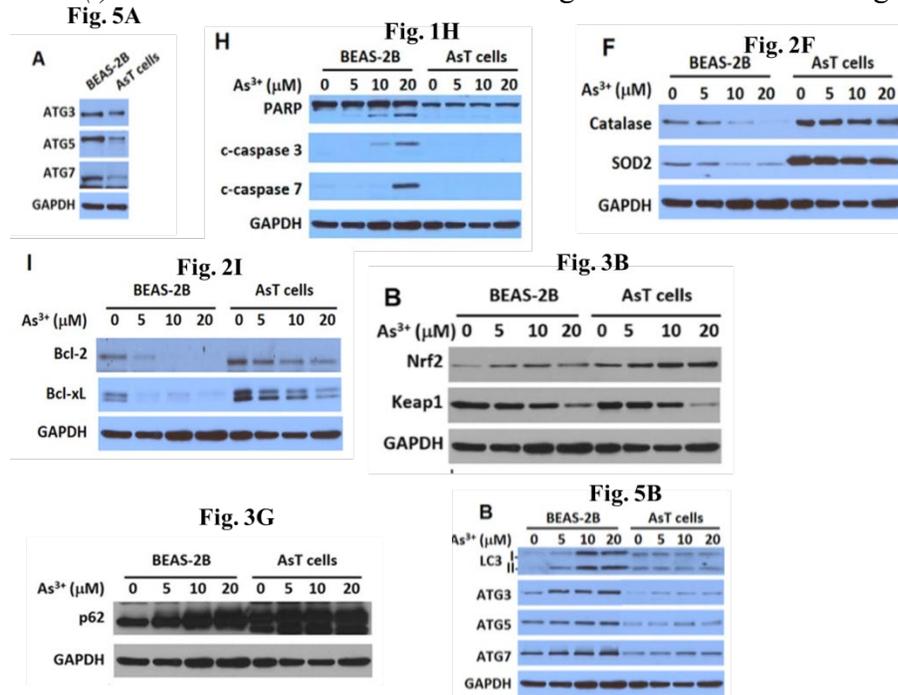
This manuscript was retracted on September 7, 2018.

10.2a. *JBC retraction notice*: The following nine issues were identified in the JBC retraction notice (Appendix 056):

- (i) The GAPDH immunoblot in Figure 5A was reused in Fig. 1H, 2F, 2I, 3B, 3G, and 5B.
- (ii) The ESR spectrum from BEAS-2B cells in Figure 2A was previously published in Son *et al.* 2014, without attribution.
- (iii) The BEAS-2B image in Figure 2C was reused in Figure 6A as control siRNA.
- (iv) The GAPDH immunoblot from Figure 3A was previously published in Figure 4A of Son *et al.* 2014, representing different experimental conditions.
- (v) The GAPDH immunoblot from Fig. 3D was reused in Fig. 3I.
- (vi) The GAPDH immunoblot was reused in Figure 4A as actin.
- (vii) Lanes 9-12 of the β -actin gel in Figure 4B was reused in lanes 9-10 of the Bcl-xL ARE F1 gel in Figure 4E.
- (viii) Figure 5B contained several undeclared splices.
- (ix) In Figure 5C, the control images for BEAS-2B and AsT cells were reused.

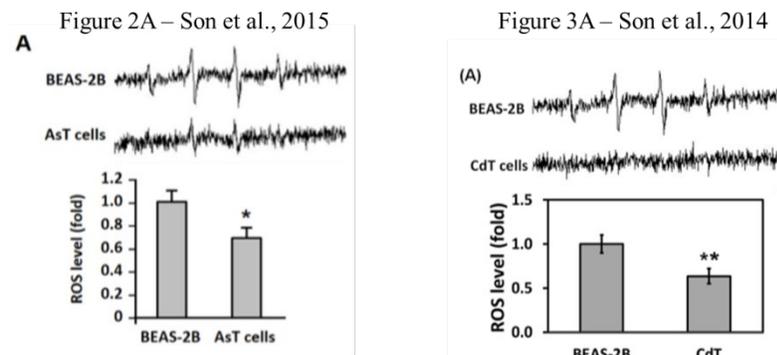
10.2b. *The committee's analysis of issues identified in the retraction notice from JBC*:

Issue (i): The GAPDH immunoblot in Figure 5A was reused in Fig. 1H, 2F, 2I, 3B, 3G, and 5B.



The committee confirmed the findings from the retraction notice. While the experimental conditions for Figs. 1H, 2F, 2I, 3B, 3G and 5B were identical, Fig. 5A presents a different experiment. Thus, Fig. 5A is a misuse of the data and the rest are inappropriate uses of load control since the same GAPDH gel was used as control for multiple blots. Since no accurate metadata were provided with the images by the Respondents, the committee could not determine when the blots were done.

Issue (ii): The ESR spectrum from BEAS-2B cells in Figure 2A was previously published in Son *et al.* 2014, without attribution.

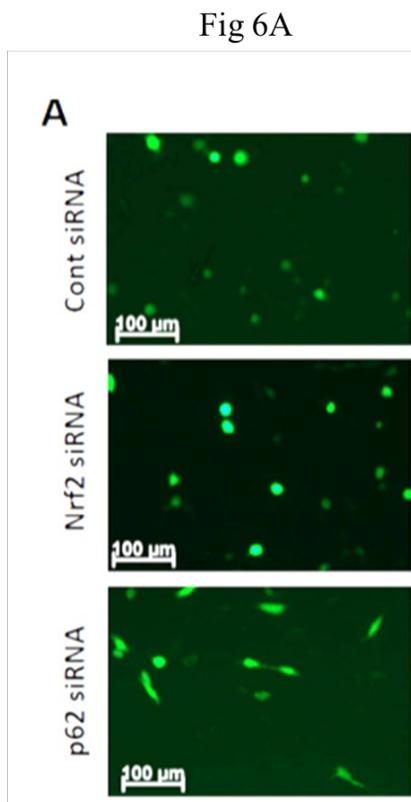
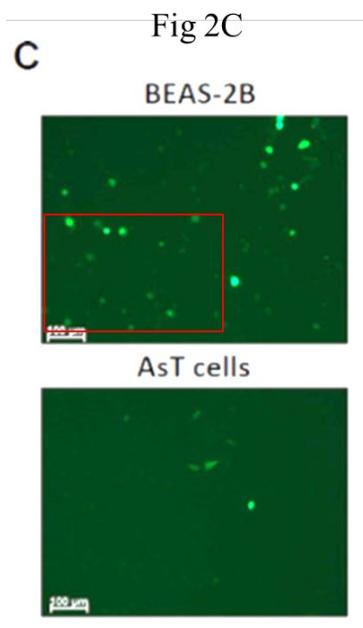


The committee concurred with the retraction notice that the ESR spectrum from BEAS-2B cells in Fig. 2A is identical to that previously published by the respondents in Fig. 3A of the Son et al., JBC, 2014, manuscript (retracted manuscript R3, Appendix 057). Since the data provided for these two figures are not raw data, the committee cannot determine if the traces from BEAS-2B cells in these two figures represent different experiments or they are from one experiment.

On a separate note, the committee is confused by the lack of consistency between ESR spectra for untreated BEAS-2B cells in the two figures above and Fig. 6A in the retracted manuscript R1 Son et al., JBC, 2017 (figure below) (Appendix 055). In Fig 6A no major peaks were noted in the untreated BEAS-2B cells, whereas in Figs. 2A and 3A discussed above there were major peaks. In fact, the controls in Figs. 2A and 3A look more like the Ni treated BEAS-2B cells in Fig. 6A.

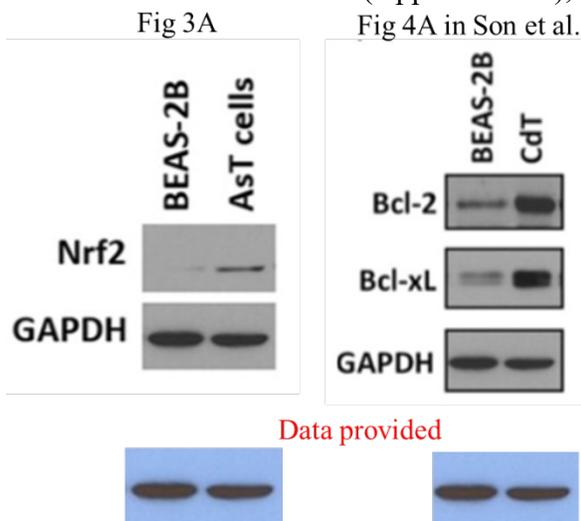


Issue (iii): The BEAS-2B image in Figure 2C was reused in Figure 6A as control siRNA.



The committee confirmed the retraction notice that the image in Fig. 2C was reused in Fig. 6A (after cropping; approximate cropping strategy is indicated by the red box) as representing a different experimental condition.

Issue (iv): The GAPDH immunoblot from Figure 3A was previously published in Figure 4A of Son *et al.* 2014 (Appendix 057), representing different experimental conditions.

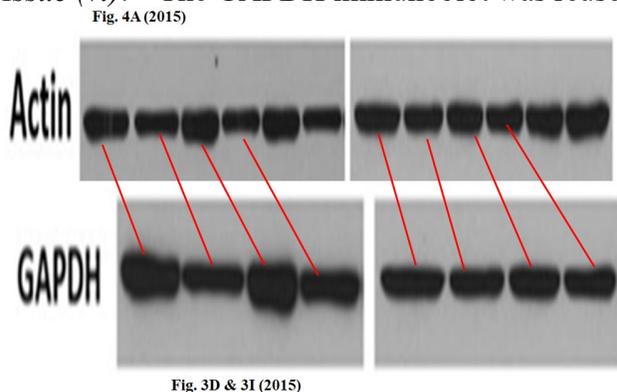


The committee concurred with the retraction statement. Even the data provided by Respondents was identical, despite the file being labeled differently. Of note, the second lanes represent

different treatment conditions in the two figures, so reusing the same load control is a misrepresentation of the data.

Issue (v): The GAPDH immunoblot from Fig. 3D was reused in Fig. 3I.
and

Issue (vi): The GAPDH immunoblot was reused in Figure 4A as actin.



The committee confirmed the retraction notice and additionally found the major inconsistencies listed in the figure and text below.

More importantly, some of the identical data are labeled differently, thus misrepresenting what they are. This **constitutes data fabrication**. **Additionally**, some of the data for the same experiments contradict each other as indicated by the dashed ovals in Figs. 3I and 4A, below. In Fig. 4A, arsenic treatment induced p62 and Nrf2 expression in all conditions (control, p62, and Nrf2-siRNA treated cells), however, in Fig. 3I, the opposite occurs in the p62-siRNA-treated cells (pink and dark purple dashed ovals). According to the figure legend and labeling these should have had similar responses.

Fig. 3D (2015)

Fig. 3I (2015)

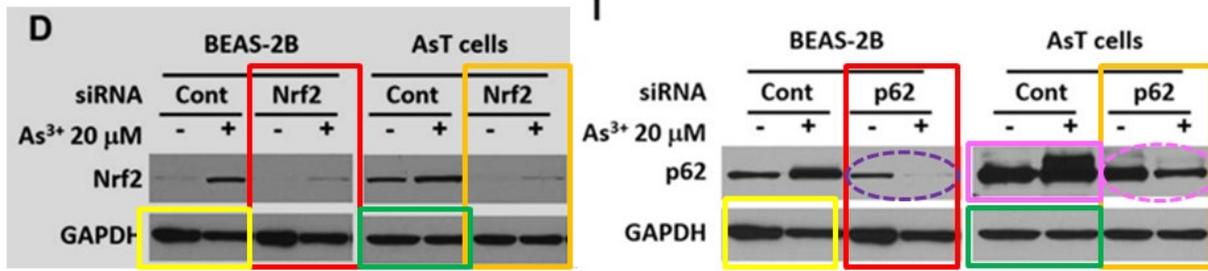
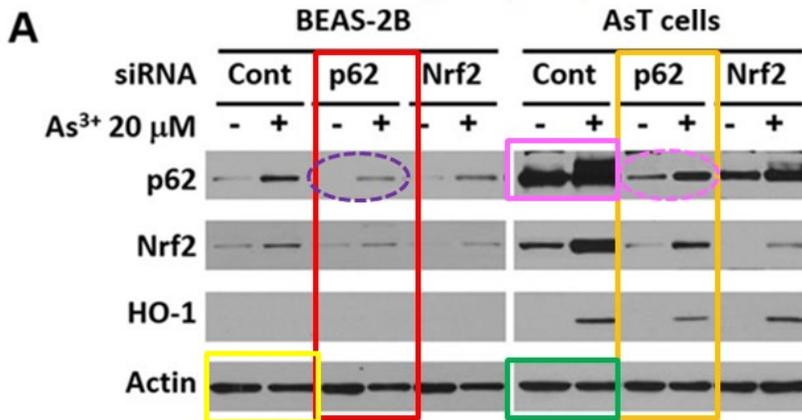


Fig. 4A (2015)



Yellow: BEAS-2B GAPDH from Fig. 3D reused in 3I and in 4A, where it is labeled as actin.

Red: BEAS-2B GAPDH from Fig. 3D reused in 3I and in 4A, where it is labeled as actin. Additionally, in Fig. 3D, the samples were labeled as from Nrf2 siRNA-treated cells whereas in Figs. 3I and 4A they are labeled as samples from p62 siRNA-treated cells.

Green: AsT cells GAPDH from Fig. 3D reused in 3I and in 4A, where it is labeled as actin.

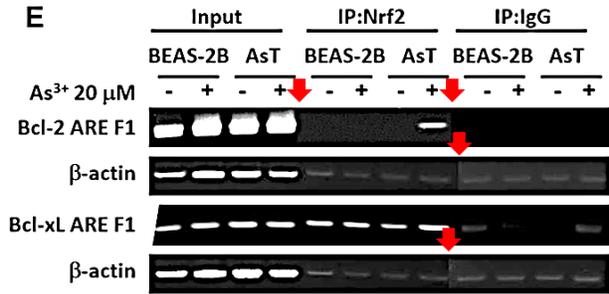
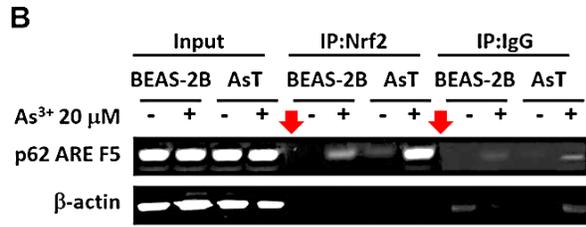
Orange: AsT cells GAPDH from Fig. 3D reused in 3I and in 4A, where it is labeled as actin. Additionally, in Fig. 3D, the samples were labeled as from Nrf2 siRNA-treated cells whereas in Figs. 3I and 4A they are labeled as samples from p62 siRNA-treated cells.

Pink Box: AsT cells p62 from Fig. 3I reused in 4A.

Pink Dashed Oval: In AsT cells treated with p62 siRNA, the data in Figs. 3I and 4A are contradictory with respect to the effect of Arsenic treatment. In Fig. 3I, p62 decreases in the presence of Arsenic, whereas in Fig. 4A, it increases.

Dark Purple Dashed Oval: In BEAS-2B cells treated with p62 siRNA, the data in Figs. 3I and 4A are contradictory with respect to the effect of Arsenic treatment. In Fig. 3I, p62 decreases in the presence of Arsenic, whereas in Fig. 4A, it increases.

Issue (vii): Lanes 9-12 of the β -actin gel in Figure 4B was reused in lanes 9-10 of the Bcl-xL ARE F1 gel in Figure 4E.



4B Actin

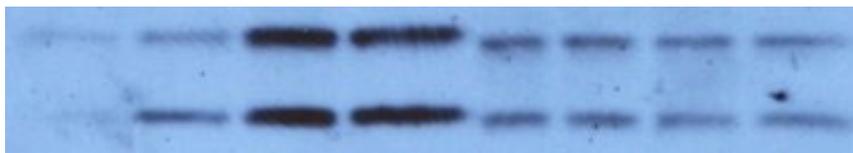


4E Bcl-xL

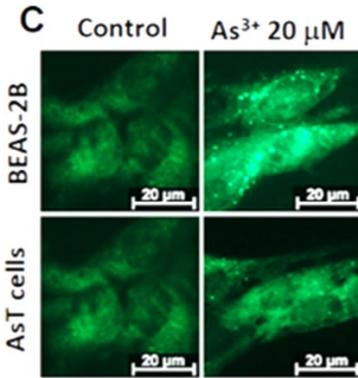
The committee concurred with the findings from JBC retraction notice regarding duplication of these lanes. **In addition** to the noted irregularity in the retraction notice, the committee also noted 6 undeclared splices as indicated by the red arrows (see above figures). The splices were confirmed in the JPG images provided by the respondents.

Issue (viii): Figure 5B contained several undeclared splices.

The committee could not find the undeclared splices in Fig. 5B as stated in the retraction notice.



Issue (ix): In Figure 5C, the control images for BEAS-2B and AsT cells were reused.



The committee confirmed that the control images for BEAS-2B and AsT cells in Fig. 5C are identical. The images provided by Respondents for this figure contain no metadata.

Overall conclusions for these nine issues: These issues represent a significant departure from accepted practices of the research community that were committed intentionally and these allegations are proven by a preponderance of evidence. The committee concluded that these are cases of data falsification and fabrication.

10.2c. *Additional issues with the manuscript identified by the committee that the JBC retraction notice did not indicate:*

10.2c1. *Figures for which no electronic data were provided:* Figs. 1G, 1I, 2D, 2G, Excel file for 2J, 2I, Excel files for 3B, 3D, 3E, 3G and 3J, Excel files for 4A, 4D and 4F, 5A, three western blot lanes in Fig. 5B, 5D, 5G, 5I, 5J, 5K, 5L, 5M, 5N, 6E, 6F, 7A, 7B.

Not preserving raw data is a significant departure from accepted practices of the research community and is a breach of NIH and UK data retention policies (Appendices 091 and 092) that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

10.2c2. *Issues identified with electronic data provided:*

Image provided for Fig. 5C, AsT cells under control conditions, does not match the published figure.

Provided by Respondents:

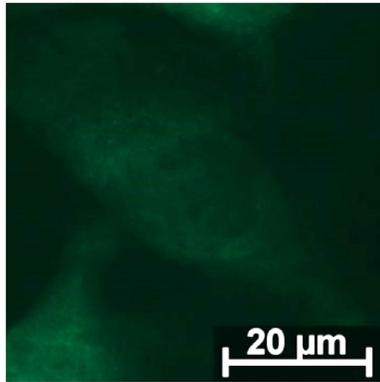
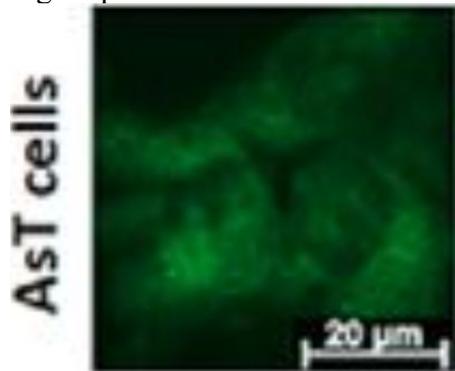


Figure published:



For Fig. 5H, some images had the date of October 2018, which is not consistent with the publication date as this indicates that the data would have been collected ~3 years after the manuscript was published.

As the Respondents could not provide the raw data used to generate Fig. 5C and provided files that were created post-publication, this is considered a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that these are cases of data fabrication.

10.2c3. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could verify:* Figs. 5B (LC3 lane); Fig 3H – p62L for AsT cells and p62S for BEAS-2B cells, only p62 written on the film; Figs. 1D, 2A, 6C, 6D.

10.2c4. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could **not** verify because of lack of labeling or because the data did not match the published figure:*

Fig. 1G: The provided data for the BEAS-2B cells does match the published data; however, the provided and published data for the AsT cells do not match.

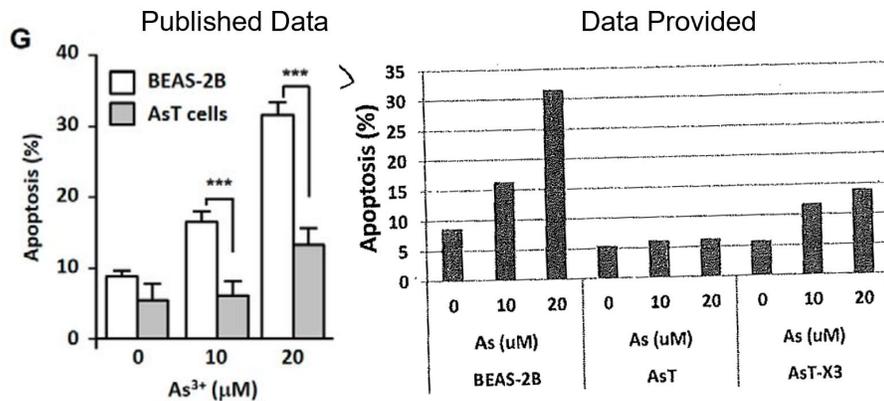


Fig. 2I: The committee could not match the Bcl-2 band from the published figure with the data on the lab notebook film. Moreover, on the raw film the concentration is marked as 0, 1, 5 and 10 whereas in the manuscript it is labeled as 0, 5, 10 and 20 μM . Additionally, the film is labelled as CrT, which likely stands for Cr-transformed cells. In the Figure AsT cells are studied since the manuscript is on the effects of As.

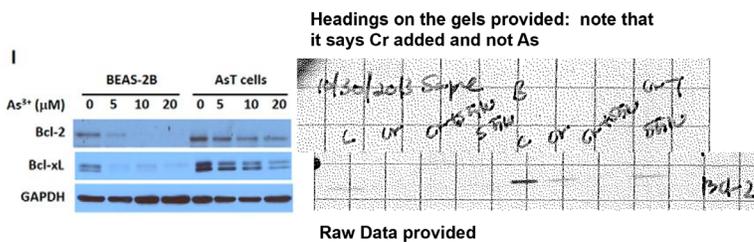


Fig. 2J: There were not labels provided on the original data provided, so it is unclear which part of the blots correspond to the data presented in the manuscript.

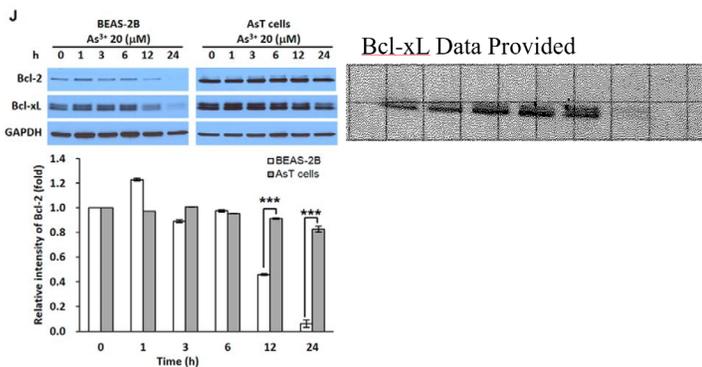


Fig. 3C: Only one of the Nrf2 blots was indicated in the provided data, and it does not match the published data.

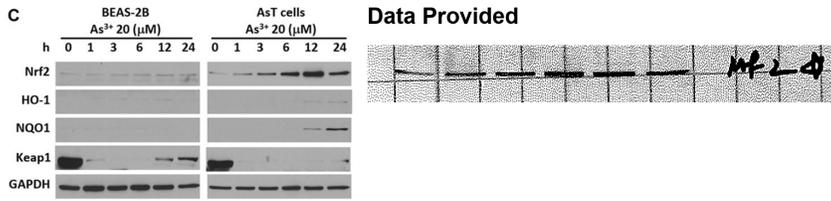


Fig. 4A: The committee could not match the data provided in lab notebook to the published figure (for both western blot, which, according to the respondents, was performed on July 26, 2014, page 16; and Excel file with the quantification, which was done on Aug 16, 2013 – page 139). The committee identified that the western blot gel was performed 11 months after the quantitation of the gels was reported in the lab notebook.

Fig. 4A

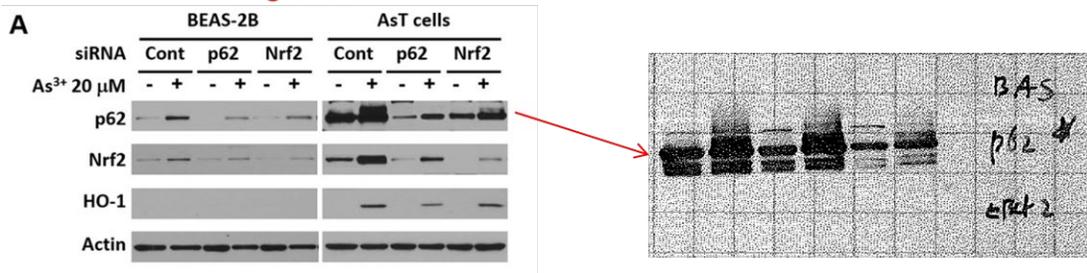


Fig. 4A: HO-1 western on the left is empty.



Fig. 5B: LC3 does not match the published figure.

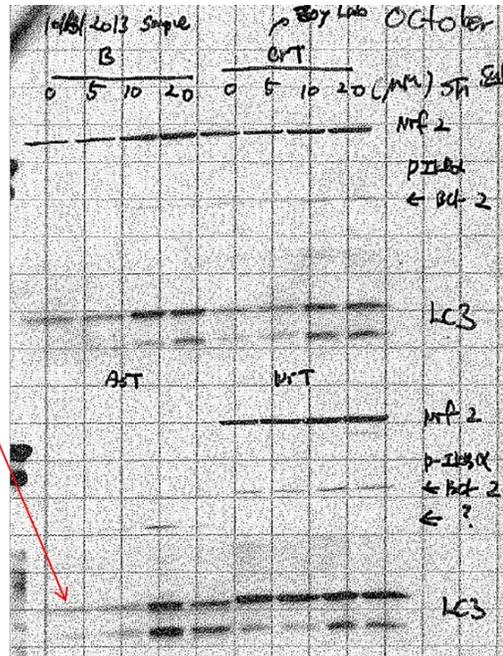
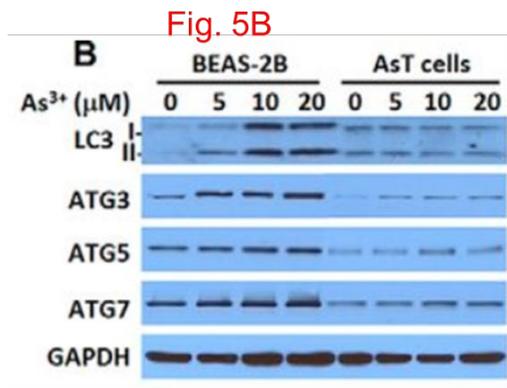


Fig. 6E: The film provided had no labels, therefore the committee could not validate the data.

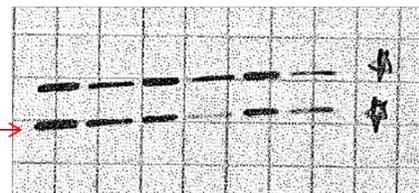
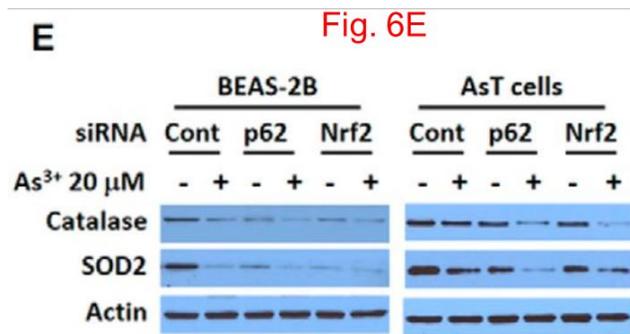
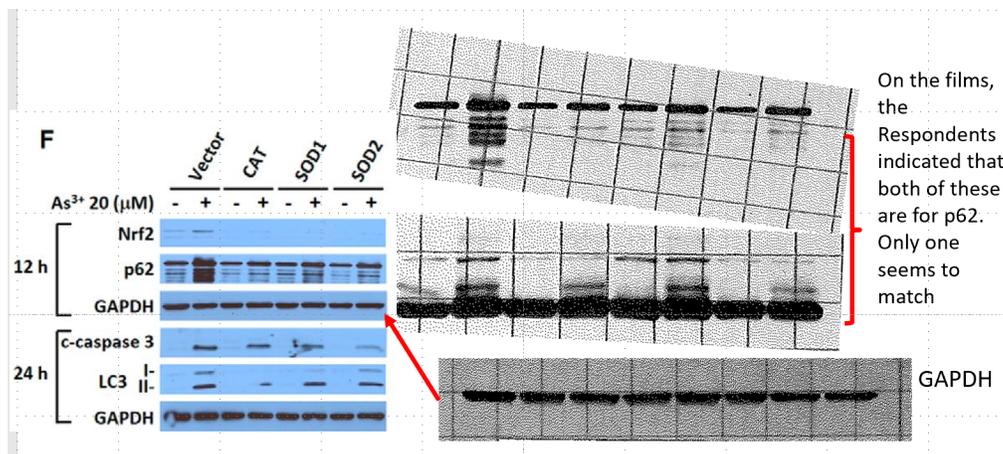


Fig. 6F: The film provided for GAPDH had no labels (dated March 6, 2014) and only matched the data for the top GAPDH blot. The bottom GAPDH blot could not be matched to the data provided. The committee could not match any of the Nrf2 data (dated March 7 or 11?). The labels on the p62 films were conflicting, and two different gels were labeled with p62. While one could represent a lesser exposure, neither were a match to the published Figure. The blots for Nrf2 were identified, and they did not match the published data either.



The raw data provided did not match the figures published, thus it was not provided accurately. Not preserving raw data used in publications is a significant departure from accepted practices of the research community that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

10.3. R3: Son, Y.O., Pratheeshkumar, P., Roy, R.V., Hitron, J.A., Wang, L., Zhang, Z., and Shi, X. Nrf2/p62 signaling in apoptosis resistance and its role in cadmium-induced carcinogenesis. *Journal of Biological Chemistry*, 293, 15455, 2014 (Appendix 057): All figures.

This manuscript was retracted on September 7, 2018.

10.3a. *JBC retraction notice*: The following nine issues were identified in the JBC retraction notice (Appendix 057):

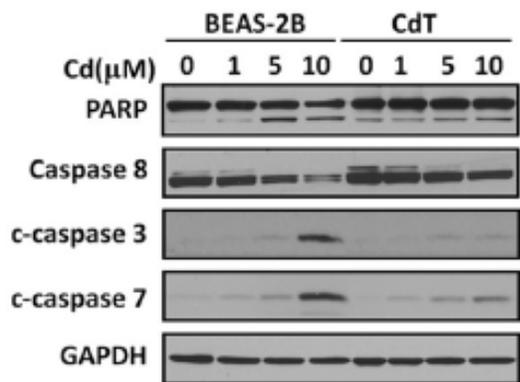
- (i) Lanes 1-3 were duplicated in lanes 5-6 of the c-caspase 3 immunoblot in Figure 2C.
- (ii) The Nrf2 immunoblot from BEAS-2B cells from Figure 5D was reused in Figure 8A.
- (iii) In Figure 8C, the first two lanes of the p62 ARE F1 and the beta-actin gels were duplicated.
- (iv) In Figure 8C, lanes 1 and 8 of the p62 ARE F4 were duplicated.
- (v) In Figure 9C, lanes 2 and 5 of the Bcl-xL ARE R1 gel were duplicated.
- (vi) In Figure 9F, lanes 1 and 2 were reused in lanes 5 and 6 of the Bcl-2 ARE F1 gel.
- (vii) In Figure 11B lanes 3 and 5 of the Nrf2 immunoblot were duplicated.
- (viii) In Figure 11B, lanes 4 and 6 of the lower bands in the LC3 immunoblot were duplicated.
- (ix) In Figure 11C, the SOD1 immunoblot was reused as the SOD2.

10.3b. *The committee's analysis of issues identified in the retraction notice from JBC*:

Issue (i): Lanes 1-3 were duplicated in lanes 5-6 of the c-caspase 3 immunoblot in Figure 2C.

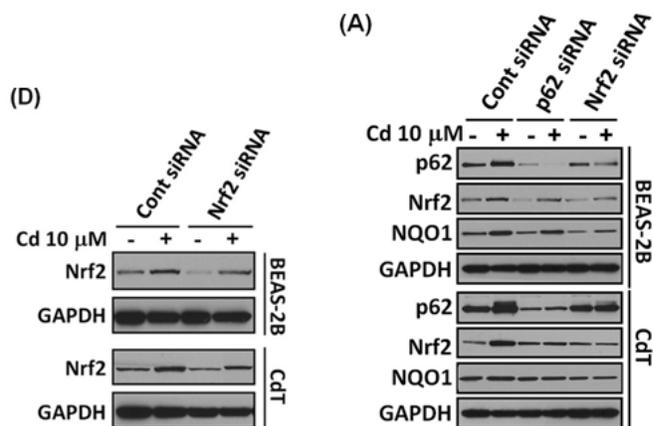


(C)



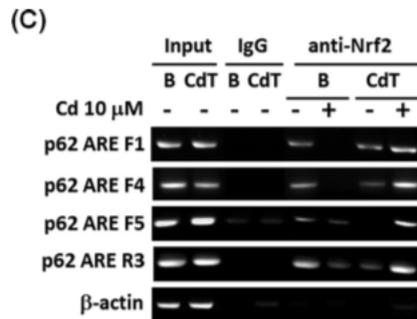
Upon examination of the data provided by the Respondents, the committee does not see how this figure was grafted together to duplicate lanes 1-3 in the 5-7 position. There are no graft lines as the committee have seen in other figures in Section 10.2b issue vii. Lanes 1 and 5 do have a similar appearance in the background dots and scratches. Lanes 3 and 7 look similar, but lanes 2 and 6 are difficult to differentiate.

Issue (ii): The Nrf2 immunoblot from BEAS-2B cells from Figure 5D was reused in Figure 8A.



For the Nrf2 immunoblot from the BEAS-2B, the first four lanes from the left towards the right of Fig. 5D were reused as the first four lanes from the left towards the right of Fig. 8A. This confirmed what was indicated in the Journal's Retraction notice. More importantly, however, in Fig. 5D the two most right lanes are labeled Nrf2 siRNA, but in Fig. 8A these two lanes are labeled p62 siRNA. This mislabeling or misrepresentation of the data changes the conclusions that can be drawn from these two Figures.

Issue (iii): In Figure 8C, the first two lanes of the p62 ARE F1 and the beta-actin gels were duplicated.

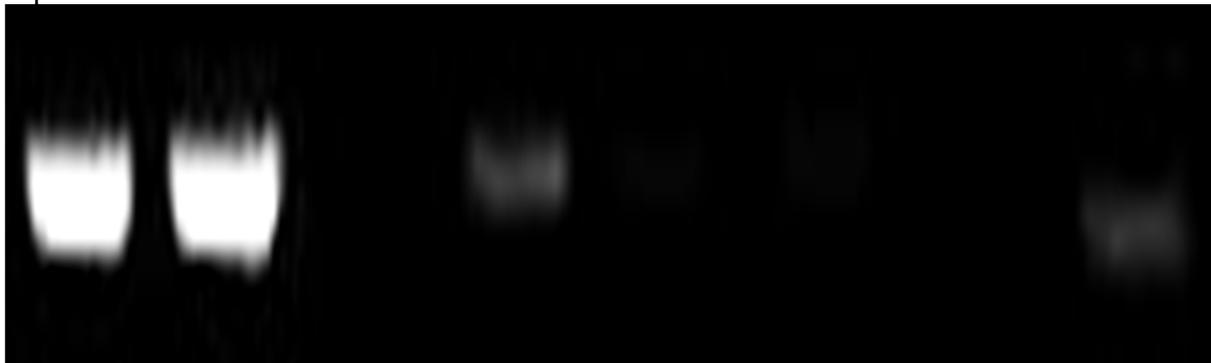


The respondents provided cropped JPG images to address this comment. Note that created dates in the metadata associated with the file are not consistent with when the experiment would have been done. Thus, the committee cannot verify that this was original data.

Data provided by the Respondents for the actin gel in Fig. 8C:



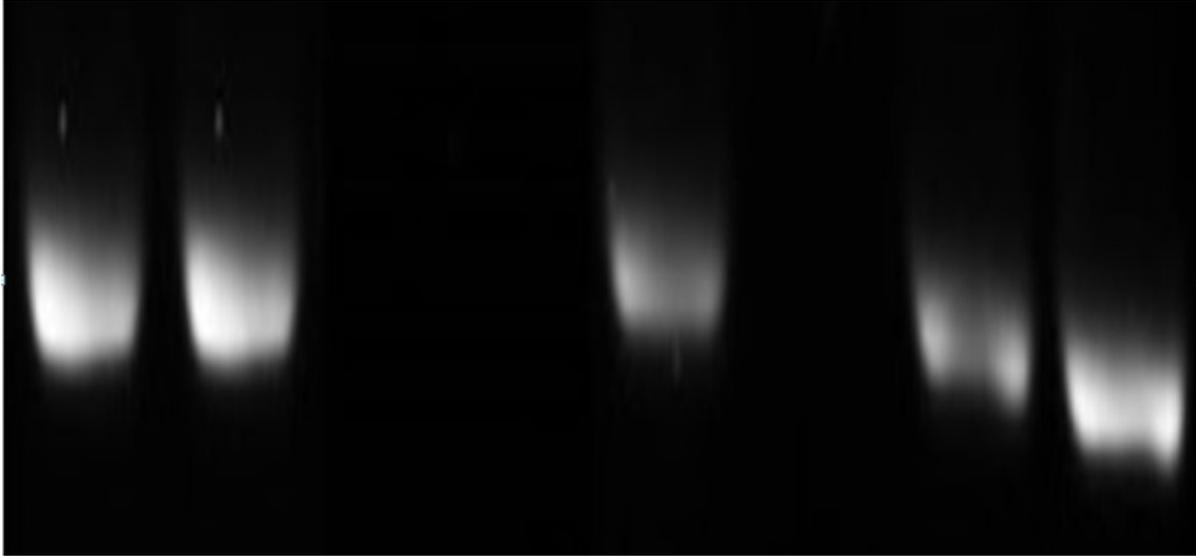
Upon examination of the data provided by the Respondents by stretching on the y-axis (below), the committee believes that the first two lanes of the beta-actin gel are different, *i.e.*, there was no duplication.



Data provided by the respondents for the p62 ARE F1 gel in Fig. 8C:



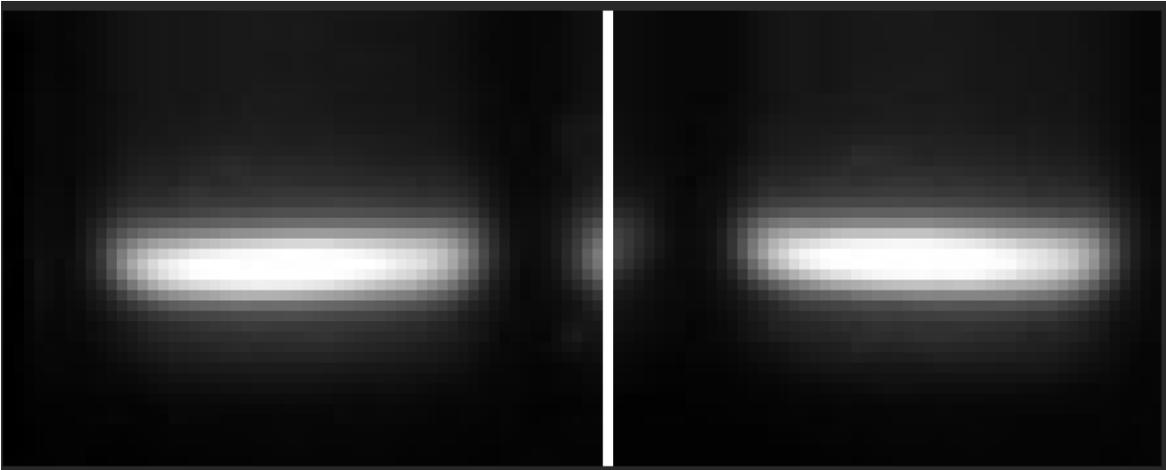
Upon examination of the data provided by the respondents by stretching on the y-axis (below), the committee believes that the first lane of the p62 ARE F1 gel was indeed duplicated in lane 2. In addition to the commentary from the retraction, the committee found that lanes 3 and 4 in this gel did not appear to have samples loaded.



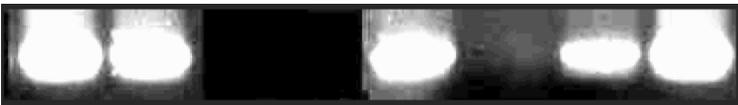
Issue (iv): In Figure 8C, lanes 1 and 8 of the p62 ARE F4 were duplicated.
Data provided by the respondents for the p62 ARE F4 gel in Fig. 8C:



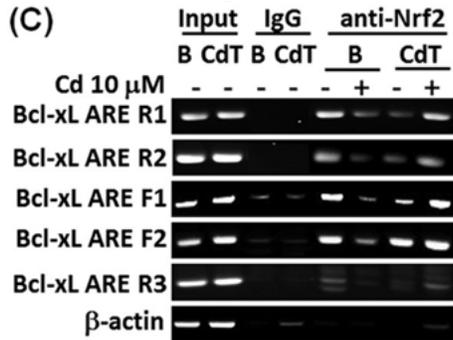
Upon examination of the data provided by the Respondents using Photoshop analysis (below), the committee confirmed that lanes 1 (left in the figure below) and 8 (right in the figure below) are identical.



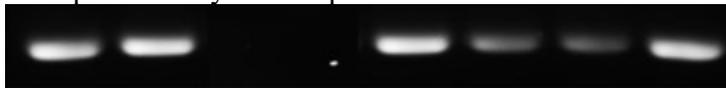
In addition to the commentary from the retraction, the committee, upon examination with Photoshop, found no pixel data in the area of the gel where lanes 3 and 4 should have been loaded. Instead, a black box was inserted in the image. It should be noted that these two lanes were meant to be a background control in the experiment. **Replacing the real gel with a black box brings into question the validity of the results and conclusions from this experiment.**



Issue (v): In Figure 9C, lanes 2 and 5 of the Bcl-xL ARE R1 gel were duplicated.



Data provided by the Respondents for the Bcl-xL ARE R1 gel in Fig. 9C:



Upon examination of the data provided by the Respondents by stretching on the y-axis (below) and Photoshop analysis, the committee confirmed that lanes 2 and 5 are identical. A clear grafting of the image can be observed, suggesting that lanes 3 and 4 are not part of the same gel. This region did appear to come from an agarose gel because signals were detected in the corresponding rectangular region of the image.

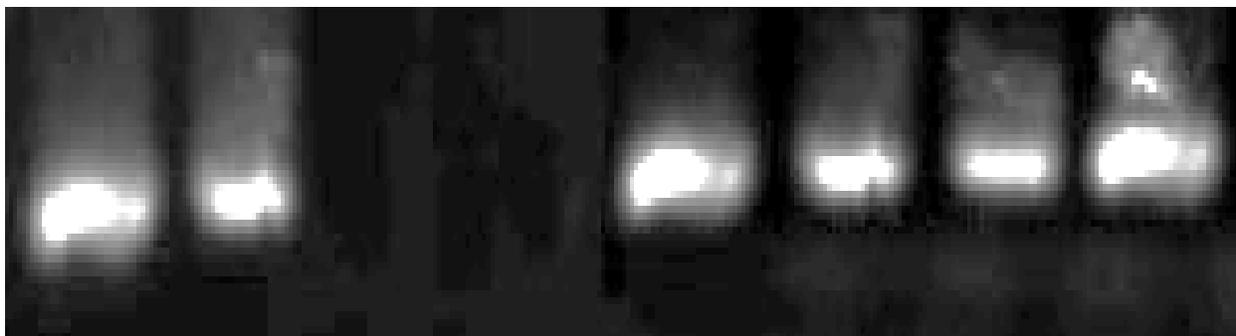


Issue (vi): In Figure 9F, lanes 1 and 2 were reused in lanes 5 and 6 of the Bcl-2 ARE F1 gel.

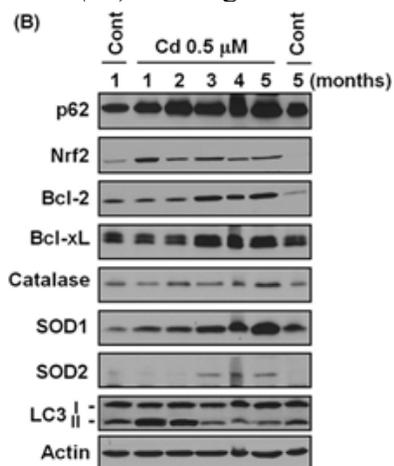
Data provided by the respondents for the Bcl-2 ARE F1 gel in Fig. 9F:



Upon examination of the data provided by the Respondents by stretching on the y-axis (below), the committee confirmed that lanes 1 and 2 were reused in lanes 5 and 6.



Issue (vii): In Figure 11B lanes 3 and 5 of the Nrf2 immunoblot were duplicated.



Data provided by the Respondents for the Nrf2 gel in Fig. 11B:



Upon examination of the data provided by the Respondents by stretching on the y-axis (below), the committee could not find any indication of grafting and splicing of the gel. However, the committee concurred that lanes 3 and 5 are the same.



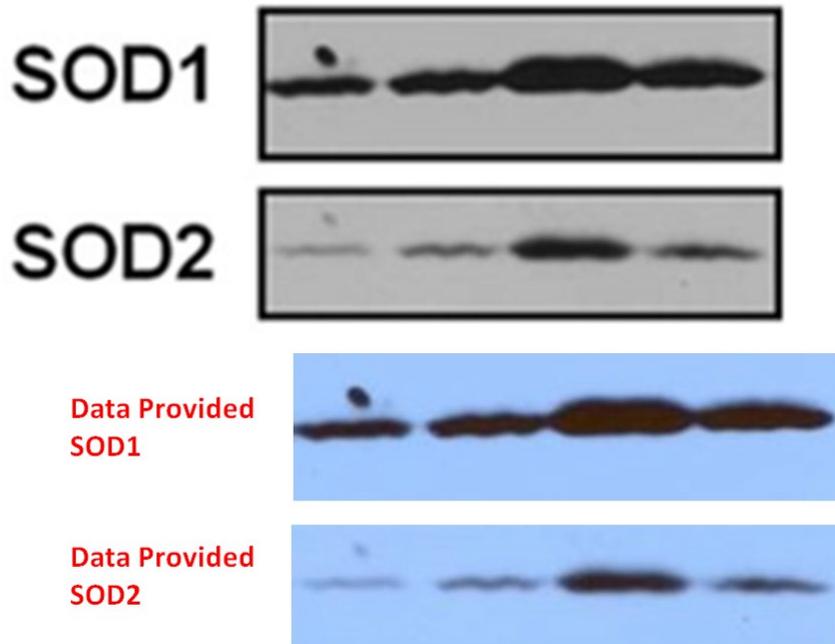
Issue (viii): In Figure 11B, lanes 4 and 6 of the lower bands in the LC3 immunoblot were duplicated.

The respondents did not provide the image for this part of Fig. 11B.

The committee's analysis could not unambiguously confirm that the lower bands in lanes 4 and 6 were duplicated. The image was cropped, magnified, and contrast was adjusted to make this assessment. The lower band is LC3II, and the upper band in the row is LC3 I. The LC3 I bands in lanes 4 and 6 are clearly different.



Issue (ix): In Figure 11C, the SOD1 immunoblot was reused as the SOD2.



The committee concurred by simple visual inspection that, as stated in the JBC retraction notice, SOD1 and SOD2 are the same blot but at different exposures. The data provided by the Respondents confirms that assessment.

The Respondents provided cropped JPG images to address this comment. Note that created dates in the metadata associated with the file are not consistent with when the experiment would have been done. Thus, the committee cannot verify that this was original data.

Additional note: The data in Figure 3 of this manuscript were used in Grant G6 Figure 8 (Appendix 011). The figure legends lacked details but were consistent.

Overall conclusions for these nine issues: These issues represent a significant departure from accepted practices of the research community that were committed intentionally and these allegations are proven by a preponderance of evidence. The committee concluded that these are cases of data falsification and fabrication.

10.3c. *Additional issues with the manuscript identified by the committee that the JBC retraction notice did not indicate:*

10.3c1. *Figures for which no electronic data were provided:* Figs. 2E, Excel file for 3A, 3D, 5D, 7B, 7D, seven western blot lanes in 8A, three western blot lanes in 10C, 10F, two panels for 10G, 10I, 10K, 11A.

Not preserving raw data is a significant departure from accepted practices of the research community and is a breach of NIH and UK data retention policies (Appendices 091 and 092) that was committed intentionally and this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

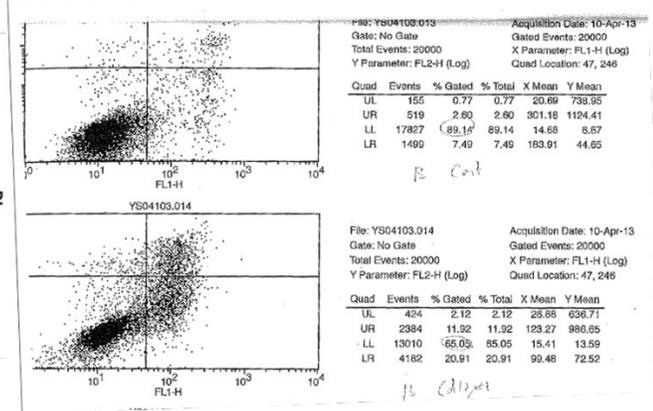
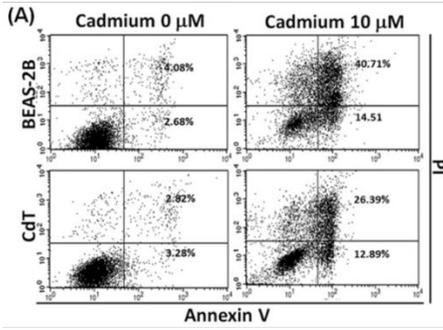
10.3c2. *Issues identified with electronic data provided:* None found.

10.3c3. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could verify:* Figs. 1C, 1D, 3G, 8D, 7D, 9D, 7C.

10.3c4. *Figures for which raw data were provided in lab notebook format by the Respondents that the committee could **not** verify because of lack of labeling or because the data did not match the published figure:*

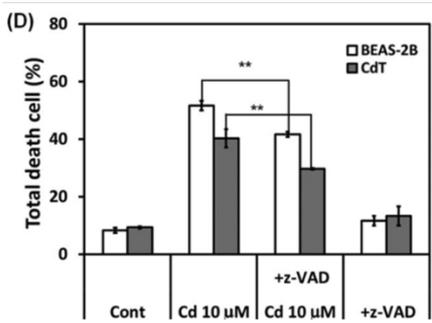
Fig. 2A: Data provided do not match the published figures.

Fig 2A



Figs. 2D and 2E: Data provided in the lab notebooks do not match the published figure; the Cd treatment data is missing in Fig. 2D. The z-VAD + Cd data is missing for both figures.

Fig 2D



February 27, 2013

Annexin V/PI

Cucurbitacin B 10nM

z-VAD 20nM

3MA 1mM

BEAS-2B	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
UR	0.85	2.91	1.85	6.67	1.57	1.73	1.57	1.73
LR	2.63	28.99	1.53	30.30	17.42	13.75	17.42	13.75
Sum	10.19	31.9	3.46	38.98	18.99	15.48	18.99	15.48

BEAS-2B	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
UR	1.51	4.35	6.46	9.15	9.93	0.75	9.93	0.75
LR	25.34	39.20	14.93	40.3	18.45	25	18.45	25
Sum	26.65	43.64	21.39	49.45	19.38	25.75	19.38	25.75

BEAS-2B	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
1	10.18	31.9	3.46	38.98	18.99	15.48	18.99	15.48
2	26.65	43.64	21.39	49.45	19.38	25.75	19.38	25.75
average	18.415	37.77	16.435	43.205	19.185	20.615	19.185	20.615

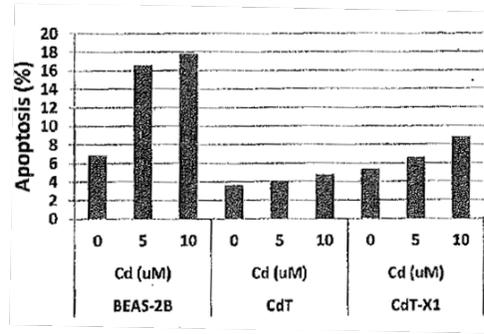
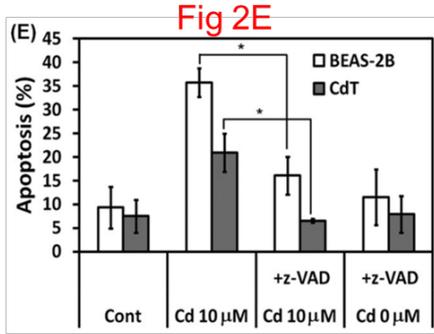
BEAS-2B cells

CdT	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
1	1.19	8.96	2	8.62	0.58	1.01	0.58	1.01
2	13.01	37.9	14.62	39.47	12.7	14.62	12.7	14.62
3	14.2	46.88	10.02	47.99	13.68	16.63	13.68	16.63

CdT-X1	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
8	0.45	2.03	1.87	3.37	0.54	0.31	0.54	0.31
1	7.87	23.89	5.55	18.61	4.12	5.95	4.12	5.95
9	8.12	22.72	7.12	22.18	4.66	6.20	4.66	6.20

CdT	Cont	CuB	CuB	CuB	+z-VAD	+3MA	+z-VAD	+3MA
6	14.2	46.88	10.02	47.99	13.68	16.63	13.68	16.63
9	8.12	22.72	7.12	22.18	4.66	6.20	4.66	6.20
5	11.40	38.29	11.87	35.085	9.17	10.948	9.17	10.948

CdT cells



February 27, 2013
Annexin V/PI
Cucurbitacin B 10μM
z-VAD 20μM
3MA 1mM

Cell Line	Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	Sum	
BEAS-2B I	0.05	2.91	1.85	6.57	1.57	1.73	
LR	8.53	28.99	7.63	30.30	17.42	13.73	
Sum	10.18	31.9	9.48	36.86	18.99	15.46	
BEAS-2B II	1.51	4.30	6.46	9.15	0.93	0.75	
LR	25.34	39.20	14.93	40.3	18.45	25	
Sum	26.85	43.64	21.39	49.45	19.38	25.75	
BEAS-2B	1	10.18	31.9	9.48	35.95	18.99	15.46
2	26.85	43.64	21.39	49.45	19.38	25.75	
average	18.415	37.77	15.435	43.205	19.185	20.605	
BEAS-2B	1	0.55	1.73	0.51	2.01	1.03	0.84
2	1.45	2.37	1.10	2.69	1.05	1.40	
average	1.00	2.05	0.84	2.35	1.04	1.12	

← BEAS-2B cells

Cell Line	Cont	CuB	CuB +z-VAD	CuB +3MA	CuB +z-VAD +3MA	Sum	
CdT	1.19	8.96	2	8.62	0.98	1.01	
Cont	13.01	37.9	14.62	39.47	12.7	14.82	
Sum	14.2	46.86	16.62	47.99	13.68	15.63	
CdT-X1	0.45	2.03	1.57	3.37	0.64	0.31	
1	7.67	23.69	5.55	18.81	4.12	5.95	
9	8.12	25.72	7.12	22.18	4.66	6.20	
CdT	6	14.2	46.86	16.62	47.99	13.68	15.63
5	11.16	36.29	11.87	35.085	9.17	10.948	
CdT	10	1.27	4.20	1.49	4.30	1.23	1.40
14	0.73	2.30	0.54	1.90	0.42	0.56	
12	1.00	3.28	1.06	3.14	0.82	0.98	

← CdT cells

Fig. 6D: Data do not match for p62. No GAPDH provided. According to the handwritten note at the top of the lab notebook page, lanes 1,2, 5 and 6 were used in the figure.

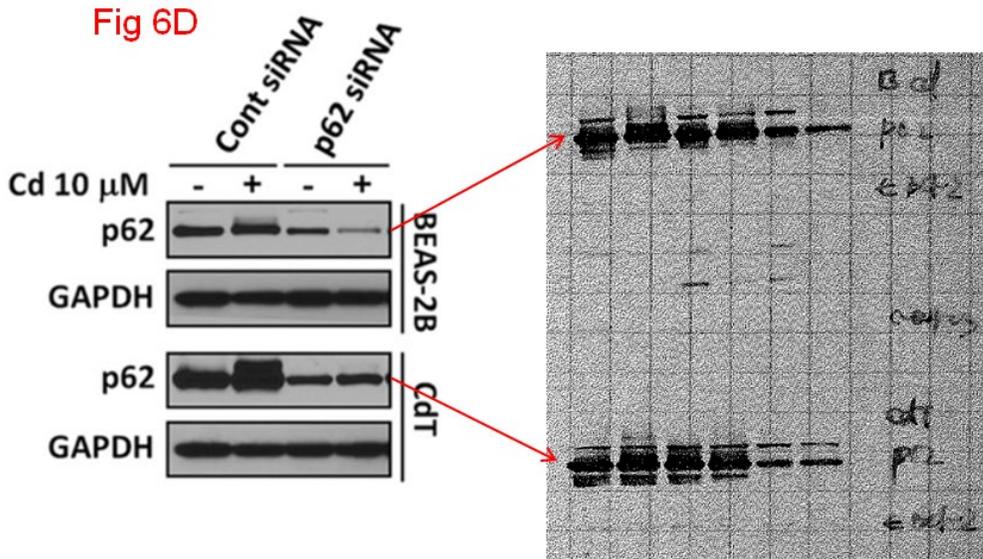


Fig. 9A: The Bcl-xL bands in the manuscript do not match those in the notebook. There are 2 bands on top of each other in the published figure but just one visible in the data from the notebook.

(A) Fig 9A

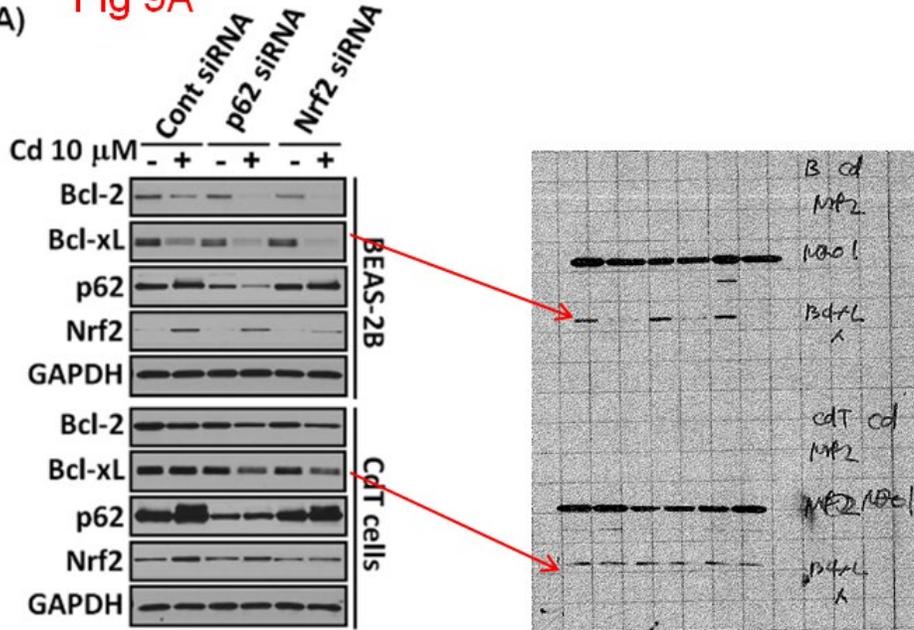
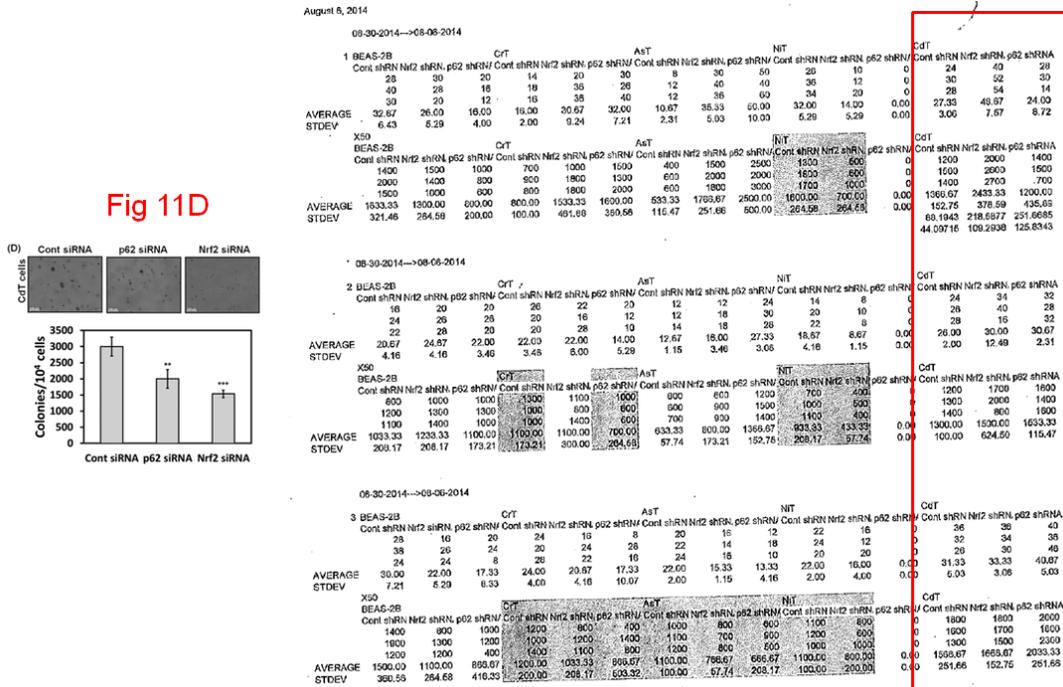


Fig. 11D: Data provided do not match the figure. Note that the figure indicates siRNA whereas the lab notebook page shows shRNA.



Since the raw data provided did not match the figures published, not preserving raw data is a significant departure from accepted practices of the research community and is a breach of NIH and UK data retention policies (Appendices 091 and 092) that was committed intentionally and

this allegation is proven by a preponderance of evidence. The committee concluded that this is a case of data fabrication.

N. RESPONDENTS' RESPONSES TO REPORT

See separate folder "Respondents responses".

O. COMMITTEE REMARKS ON RESPONDENTS' COMMENTS

The Respondents have noted that the manipulations this group did in creating the figures highlighted by the committee's report did not constitute misconduct or falsification because "they did not change the conclusions of the experiment". The committee used the definition for research misconduct as posted by the Federal Office of Research Integrity within the U.S. Department of Health and Human Services (<https://ori.hhs.gov/definition-misconduct>), which concepts are reiterated in the University's Administrative Regulation on research misconduct:

Research misconduct means fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.

(a) Fabrication is making up data or results and recording or reporting them.

*(b) Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results **such that the research is not accurately represented in the research record.** (Committee's emphasis)*

(c) Plagiarism is the appropriation of another person's ideas, processes, results, or words without giving appropriate credit.

(d) Research misconduct does not include honest error or differences of opinion.

The committee extensively discussed the role of "intent" in the actions of the respondents. To better define its role, the committee drew from a report to the Secretary of HHS made by a select commission in 1995 and excerpted below. The fact that the committee found so many cases where the respondents published data that were not accurately presented was interpreted as reckless, in keeping with the Commission's recommendations.

*"An intent to deceive is often difficult to prove; proof almost always relies on circumstantial evidence, which can, however, include an analysis of the behavior of the person accused of misconduct. One commonly accepted principle, adopted by the Commission, is that an intent to deceive may be inferred from a person's acting **in reckless disregard for the truth** (committee's emphasis). Conduct that is merely careless or inadvertent is not included in the Commission's proposed definition of research misconduct. However, the Commission intends that such careless conduct continue to be addressed in the high standards of grant application review and in institutional and professional standards for appointment, promotion, publication, and other incentives."* (https://ori.hhs.gov/sites/default/files/report_commission.pdf)

The AR 7.1 defines Research Misconduct as:

"A finding of research misconduct requires that the events constitute research misconduct as defined in Section II, above, and that:

1. There is a significant departure from accepted practices of the relevant research community; and
2. The misconduct is committed intentionally, or knowingly, or **recklessly** (*committee's emphasis*); and
3. The allegation is proven by a preponderance of evidence”

The Respondents' defense highlights a greater problem with the work published by this group and a cavalier attitude about figure construction. Data published in the literature comes with certain expectations. The reader expects that published experimental data are presented accurately and the experiments are performed as described. This allows the reader to make their own interpretations of the results and determine if he/she agrees with the authors. Such a critical reading of published literature is at the very heart of the scientific process. However, there is an additional way a reader uses published experiment data. He/she may use it to interpret their own experimental results. To do this, a reader expects that the published data is accurately presented so that valid comparisons to other data can be made. A published figure may inform about cell morphology (size and shape), the migration of a protein on SDS-PAGE, or the relative abundance of a protein on western blot. This valuable information is not necessarily relevant to the conclusions made by the original authors. If the published data are altered, then the reader may unknowingly make an incorrect conclusion about his/her own data. It is this aspect of the scientific process that has been thwarted by the Respondents' actions.

Specific Responses to the Rebuttal Comments:

Point 1a: The committee interviewed Dr. DiPaola extensively (twice) about his role in the work published in the Wang et al. paper and in the relinquished grant (see transcripts in Appendices 048 and 060).

Dr. Wang's interview about his interactions with Dr. DiPaola (Appendix 058).

a. Appendix 058, page 7: "... I switched to Dr. DiPaola's lab. So I worked as a scientist in his lab. But still sitting in the same office, sitting at the same bench."

b. Appendix 058, page 30:

Committee: "Who did you meet with more when you were putting the paper together?"

Dr. Wang: "Of course, it is Dr. Shi. Yeah, his office is close to me. *I sit in his lab* [committee's emphasis], and, so, I meet him more. And Dr. Zhang is also more. She is close to me. And Dr DiPaola is weekly or biweekly meeting together but I am not always joining the meeting."

c. Appendix 058, page 28, regarding the Wang et al. paper: "The writing process, the draft is by me and Dr. Shi."

d. Appendix 058, page 8, Wang: "They [*assumed to be Shi and Zhang*] just told me the original data you should keep by yourself"

Quotes from the 1st interview with Dr. DiPaola (Appendix 048):

Appendix 048, page 10:

Committee: "So did you interact with any other lab members, other than Dr. Shi and Dr. Zhang?"

Dr. DiPaola: "They will bring in Dr. Lei Wang periodically, the first author."

Appendix 048, page 12-13: “You know, they drafted this. And so, you know, I was really, you know, given a very completed draft with, you know, what we have been discussing in terms of, you know, they put all the figures in. I got to, you know, of course, edit it. And most of the edits were, you know, were shaping the language, you know, and so forth”

Appendix 048, page 13: “I’ll be honest, I’m not sure why I ended up being the corresponding, you know, author on this. As best as I can recall I kind of remember them somewhat insisting. I thought they were just being, you know, -- you know, I don’t know, helpful, that I was, you know, an expert in prostate cancer.”

Interview with Dr. Shi (Appendix 052)

Appendix 052, page 34, Shi “So did he (Dr. DiPaola) see the original data?

He see some of it”

Appendix 052, page 31, Shi: “I got a lot of kind of input from him (DiPaola) in a clinical aspect, the importance. For example, I suggest credit for your (DiPaola’s) work. And he said, No, that’s too far from clinical, so I had to modify that.”

Appendix 052, page 32, Shi: “And sometimes Dr. Lei, Dr. Wang, give the result to me, I give it to Dr. DiPaola.”

Based on these interviews, the committee determined that Dr. DiPaola’s oversight of the experiments and of the preparation of the figures and text was minimal. He was mainly shown PowerPoint compilations of the data, instead of seeing the raw data for himself. He relied on the judgement of Dr. Shi in evaluating the veracity of the data presented and its meaning. Furthermore, the suggestion that Dr. DiPaola serve as corresponding author was made by Dr. Shi. From his Dean’s accounts, Dr. DiPaola did provide salary funding for the postdoc doing part of the experiments but did not direct the day-to-day actions of the post-doc (Dr. Wang) who continued to work in Dr. Shi’s laboratory (this was confirmed in the interviews by Drs. Zhang, DiPaola, and Wang). Dr. DiPaola also provided funds to support the research (this was confirmed in the interviews of Drs. Shi, Zhang, and DiPaola). Despite assuming the senior author position on the publication, the committee felt that Dr. DiPaola had very little role in supervising the work being done and the writing of the manuscript and construction of the figures (this was confirmed in the interviews with Dr. Wang, Dr. Kim, and Dr. DiPaola). In fact, the committee questioned Dr. DiPaola’s reasons for being senior author and concluded that Dr. Shi’s insistence was the major factor. Moreover, the biggest issues in the Wang *et al.* paper are with experiments and figures generated by Dr. Kim, who was supervised directly by Dr. Zhang and who did not present his data to Dr. DiPaola (or Dr. Wang, the first author). For these reasons, Dr. DiPaola was not included as a respondent.

Point 1b: Dr. Whiteheart was appointed the Director of the MD/PhD Program in June 2019. Dr. Whiteheart’s candidacy for the directorship was instigated by the MD/PhD students and the staff of the program. His appointment was made after an internal search and interviews with several (at least 4) candidates. The selection committee recommended Dr. Whiteheart for the position and named two other individuals as Co-Directors to assist Dr. Whiteheart. This plan was first approved by Dr. Griffith, the Vice Dean for Education, and then by Dr. DiPaola. Dr. Griffith was responsible

for the negotiations. This process and appointment are in no way related to the misconduct investigation of the respondents and cannot logically be considered as a conflict of interest.

Point 2: Given the volume of instances of data irregularities noted by the committee, and the fact that multiple people were involved, the committee found it difficult to assign blame to specific individuals for each of the many irregularities. However, from the interviews it was clear that Drs. Shi and Zhang had created an environment in their laboratory that did not promote scientifically rigorous and true representations of the raw data in grants and publications. It is this behavior that the committee considers reckless. Dr. Kim was complicit in some of these incidents.

Andrew Hitron, a graduate student in Dr. Shi's laboratory, was interviewed regarding general laboratory practices. When asked about the lab procedures for recording and saving experimental protocols, and results and data, Hitron responded that it was "all up to the individuals." He agreed with the statement that there was no laboratory training program or requirement for how data should be preserved. (Appendix 051, pg. 9) Lei Wang stated that he kept original films in some of his notebooks, and the rest of the films in his lab, but because of changes to the laboratory, "I am not sure that I can find all of them." (Appendix 058, pg. 16) Hitron stated that many others in the laboratory did not annotate their films and that while searching for original data as per the committee requests, he found many films that had no labels. As a result, he "found a lot of x-ray films that just I don't know what they were supposed to be." (Appendix 051, pg. 11) Kim said that he would scan the original films of blots to his computer and store the actual films in "my drawer." Kim further noted that he would not put detailed labels on the films, "just date here and just time right there." (Appendix 049, pg. 29) He further stated that Zhang never asked what the original films looked like. (Appendix 049, pp. 44-45) These statements by employees in the Respondents' laboratories indicate that the employees in those laboratories were never given clear and consistent instructions about how to label and store raw data, leading to issues with the representations of that data in grants and publications.

Point 3: The point of the interviews was not to have the respondents defend every detail of the many problematic figures identified by the committee, but to provide representative examples of the types of issues that the committee found. Using this strategy, the committee could gain an understanding of how the respondents felt about preserving, handling and manipulating the experimental data and about ensuring that the published figures are scientifically rigorous and true representations of the raw data.

Point 5: The committee notes that during the interview, Dr. Shi indicated that "our laboratory does not have any very sophisticated kind of a technique. For example, western is 90 percent." (Appendix 052, page 9) From the observations of the committee, Drs. Shi's and Zhang's laboratories use films for western blot experiments, and the developed film constitutes the raw data. Therefore, the committee's repeated requests for hardcopy data was fully justified. The committee believes it could rightfully assert, given the nature of the research, and its review of the processes of the labs, and in line with the University's Administrative Regulation expectation that the committee can evaluate whether "[t]here is a significant departure from accepted practices of

the relevant research committee” that the failure of the respondents to maintain what was clearly raw data was a significant departure from such accepted practices, done, at the very least in many instances, recklessly. The respondents also state “Appendix 91 is the state of Kentucky and not UK regulation.” They ignore the obvious fact that page 2 of Appendix 91 clearly states that the “State University Model Records Retention Schedule” applies to all university records, and that these records can only be disposed of with the approval of the Commission. The same page also references that documents shall be kept whatever the medium (electronic or hard copy). The University is a part of the state and that which the state is obligated to do, the University is similarly obligated to do unless there is some stated exception.

Further, beyond the state policy that applies to the University, the University has maintained “Data Retention & Ownership Policy” for many years, including as last updated as of January 25, 2018 (updating a previous version that existed from before the Respondents joined the University that included the same expectation of a minimum five year retention period for all data), which is published through a variety of University websites, including the University’s ORI pages on “Policies & Guidance” (<https://www.research.uky.edu/office-research-integrity/university-kentucky-data-retention-ownership-policy>) and linked to from other University pages such as the UK Libraries page on research data services (<https://libguides.uky.edu/research data/UKYsupport>) (See attached Appendix 139). Whether the respondents received specific training on the University, state or federal requirements does not absolve them or any other University researcher from the obligations. The respondents’ reference to the loss of raw data “as a lack of transparency and is not data fabrication or falsification” is a misrepresentation of the Federal ORI reference to which they cite. The Federal ORI reference lists a failure to produce raw data as a sign of lack of transparency and therefore a red flag sign of research misconduct.

Point 6: Drs. Despa and Whiteheart have widely published cell biology studies that involve the methods and concepts covered in the publications and grants analyzed. Dr. Garneau-Tsodikova’s expertise in chemical biology further solidifies the committee’s ability to evaluate the respondent’s work.

Point 8: Because of the high frequency at which issues were detected, the committee disagrees that these could possibly be considered “honest errors”. Based on the pattern of issues over many years and the number of lab personnel involved, the committee concludes that there was a systematic disregard for accurate data presentation by the respondents.

Point 12: In fact, the committee was able to find one of the files, and this is what allowed it to conclude that the slides sent to the committee by Dr. Kim on 2/21/19 (object of section 9.1) were fabricated. The hard drive of Dr. Kim’s computer and the Share Drive used by the labs of Drs. Shi and Zhang were sequestered at the beginning of this investigation (on 6/7/18 and 6/22/18, respectively). Thus, the committee had access to all the files created/modified up to that date. Regarding issues in sections 9.1 and 9.2, the committee followed the filepaths indicated by Dr. Kim to search for the files from which Dr. Kim excerpted the slides/pages that he (directly or through Dr. Shi) sent to the committee (Appendices 135 and 137). The committee **found** the file

for section 9.1 on the sequestered hard drives, and this file was last updated/saved on 3/10/16 (5/20/18 for the same file located on Kim's computer rather than the shared drive). However, the two slides sent to the committee by Dr. Kim on 2/21/19 **did not exist** in this file (i.e., the file found on the hard drive sequestered at the beginning of the investigation following the filepath provided by Dr. Kim). Therefore, the committee can only conclude that these two slides were inserted into the file at a later date, after the start of the investigation (and many months after the experiment was performed). The files indicated by Dr. Kim for section 9.2 **do not exist** on the sequestered hard drives. Consequently, the committee concluded that the material sent by Dr. Kim to the committee was fabricated and is the topic of section 9.1 of the report.

Appendix 35, the DVD of Dr. Kim's demonstration of the use of the laboratory microscope on October 19, 2018, can be ordered from AN/DOR Reporting and Video Techniques, 179 East Maxwell Street, Lexington, KY 40508. Their phone number 859-254-0568 and their email contact information is setdepovideo@andorreporting.com. The title of the DVD provided to the committee was "Kim Investigation, October 19, 2018."

Point 13: Dr. Whiteheart was confused because he saw that there was no way the respondents could be concluding what they did from the data they showed. Dr. Whiteheart sought to move the interview to the next step instead of belaboring an obviously untenable assertion by the interviewee.

Point 18: Although not the corresponding author, Dr. Shi was nonetheless the last author on a retracted JBC 2002 publication. This listing position generally implies a leadership role in the publication.

Point 20: The letters of support for the respondents generally confirm the data-presentation irregularities noted by the committee, although the letter writers did not agree with the committee's conclusions of misconduct. They did not cite NIH guidelines, so their comments and judgements are only marginally useful. Two letter writers chose not to use their University letterhead.

Clarifications to the initial report:

On page 120, bottom paragraph; the committee incorrectly identified Dr. Shi as the corresponding author of a retracted 2002 *J. Biol. Chem.* publication when in fact he was the last author.

In section 9 (starting on page 117) it was not that the committee could not find the files. The files on the sequestered hard drives either did not exist or did not contain the slides provided to the committee by the Respondents.

P. RECOMMENDATIONS

- Given the volume and nature of the problems the committee identified in the work of Drs. Shi Zhang, and Kim, the committee recommends that disciplinary action be taken, including, but not necessarily limited to:

- 1) The Respondents should not be allowed to directly supervise any trainees (visiting scientist, postdoctorate student, or graduate student) and other laboratory personnel until they have demonstrated that they have corrected their data storage system and processing, and developed a plan for training existing and future lab personnel in the proper conduct of research. The VPR should appoint a panel to approve the Respondents' corrective actions and monitor their compliance.
- 2) Future publications and grant submissions from the Respondents should be vetted by examination of all original data (not PowerPoint) by a panel of University colleagues (appointed by the VPR) prior to submission. This should be continued at the discretion of the VPR.
- 3) The Respondents should correct or retract the publications identified in this report as containing data that could not be validated with original data, or as falsified or fabricated. The Respondents should document and provide to the VPR proof of all their correspondence with each journal involved. The correspondence with the journal editors should be consistent with the findings and documentation of this report.
- 4) The Respondents and all laboratory personnel must take an in person Responsible Conduct of Research class (*e.g.*, TOX 600, PHS 711-001) and their attendance must be verified.
- 5) The granting agencies which sponsored the research reported in the articles which contain falsified data should be notified.
- 6) The granting agencies which received applications containing fabricated or falsified data should be notified.
- 7) The Respondents should be subject to institutional actions as outlined in AR 7.1-XIII.B (the University's Research Misconduct AR provisions on institutional actions and to whom the outcome of the investigation may be communicated).

- Due to the large volume of material, the committee was only able to examine a representative sample of the publications and grants from Drs. Zhang and Shi's group. However, given the high frequency of issues detected in that representative sample, the committee recommends that further, more extensive examination of all of the Respondents' work should be considered.

- Notwithstanding that direct oversight of Dr. Wang was found not to be within Dr. DiPaola's purview, the committee feels that Dr. DiPaola failed to adequately assure himself that the research and presentation of data and findings in the Wang *et al.* (*Prostate*, 2018) paper was conducted as properly as he could reasonably control. He should retract the Wang *et al.* paper in *Prostate*, which contained many of the issues noted by the committee.