



RESEARCH ARTICLE

**REVISED** Results of an attempt to reproduce the STAP phenomenon [version 2; referees: 2 approved]

Shinichi Aizawa

Scientific Validity Examination Team, RIKEN, Kobe, Japan

**v2** First published: 01 Jun 2016, 5:1056 (doi: [10.12688/f1000research.8731.1](https://doi.org/10.12688/f1000research.8731.1))  
 Latest published: 17 Oct 2016, 5:1056 (doi: [10.12688/f1000research.8731.2](https://doi.org/10.12688/f1000research.8731.2))

**Abstract**

In 2014, Obokata and colleagues reported their observation of a novel cell reprogramming phenomenon they named ‘stimulus-triggered acquisition of pluripotency’ (STAP). The most conclusive evidence for the pluripotency of so-called STAP cells was the purported ability of such cells to contribute to chimera formation. Here, I report the results of an attempt by Haruko Obokata to replicate the phenomenon under the supervision of the Scientific Validity Examination Team of RIKEN. In this follow-up study, putative STAP cells prepared by Haruko Obokata were injected into 1051 embryos, of which 591 were recovered. However, the injected cells made no significant contribution in any of the embryos that developed.



This article is included in the **Preclinical Reproducibility and Robustness** channel.

**Open Peer Review**

Referee Status:

	Invited Referees	
	1	2
<b>REVISED</b>		
<b>version 2</b> published 17 Oct 2016		
<b>version 1</b> published 01 Jun 2016	 report	 report

- 1 **Austin Smith**, University of Cambridge UK
- 2 **Irene de Lazaro**, The University of Manchester UK

**Discuss this article**

Comments (0)

**Corresponding author:** Shinichi Aizawa ([saizawa@cdb.riken.jp](mailto:saizawa@cdb.riken.jp))

**How to cite this article:** Aizawa S. Results of an attempt to reproduce the STAP phenomenon [version 2; referees: 2 approved] *F1000Research* 2016, 5:1056 (doi: [10.12688/f1000research.8731.2](https://doi.org/10.12688/f1000research.8731.2))

**Copyright:** © 2016 Aizawa S. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Grant information:** This examination was supported by a grant for Scientific Validity Examination by RIKEN President’s office. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Competing interests:** No competing interests were disclosed.

**First published:** 01 Jun 2016, 5:1056 (doi: [10.12688/f1000research.8731.1](https://doi.org/10.12688/f1000research.8731.1))

**REVISED Amendments from Version 1**

I thank both reviewers for their comments. My responses are as follows:

- 1) We have now included a statistical analysis (t-test), which indicates that the efficiency of cell aggregate formation is significantly different between ATP treatment and HCl treatment in the C57BL/6 background. However, the difference is slight. We have revised the manuscript accordingly (Table 1 and page 5 in the text).
- 2) The effects on both cell aggregate formation and chimeric potency of the spleens' genetic background were examined in the C57BL/6 and F1(C57BL6 x 129) background. It is well known that ES culture is strongly influenced by genetic background. Both of these backgrounds were used in the retracted Nature papers. I have now revised the manuscript (page 4 and page 6) to clarify this point.
- 3) The two reports (Tang *et al.*, 2014, De los Angeles *et al.*, 2015) are now cited and briefly discussed (page 8–9).
- 4) The headings in Table 1 have been changed as suggested.
- 5) The retracted Nature papers show extensive colonization of the cells. The article reported eight chimeric embryos, showing more than 50% coat color contribution, of 48 chimeras obtained; these animals yielded "STAP"-derived offspring. This was the central finding in the now-retracted STAP reports. In the present study, no chimera equivalent to those in the retracted Nature papers were obtained, nor were any chimeras obtained showing more than 50% contribution to coat pigmentation. Indeed, no chimera showing more than 1% contribution was obtained. I have revised the text accordingly to indicate this.

See referee reports

## Introduction

Induced pluripotent stem cells (iPSCs), first reported by Takahashi and Yamanaka using a combination of exogenous genetic factors, have transformed our understanding of the gene regulatory mechanisms underlying cellular pluripotency and differentiation (Takahashi & Yamanaka, 2006). This discovery raised the possibility that cellular reprogramming may also be induced by activating endogenous pluripotency genes under certain conditions. In two reports published in *Nature* by Obokata *et al.* (2014a); Obokata *et al.* (2014b), the authors claimed to have observed how "external stimuli such as a transient low-pH stressor reprogram somatic cells into pluripotent cells," which they referred to as the STAP phenomenon; subsequently, however, after multiple problems were found with the handling and presentation of the data in a pattern indicative of research misconduct, both papers were retracted.

The present article reports the results of a study conducted by Haruko Obokata in the RIKEN Center for Developmental Biology (CDB), which was designed to determine whether the STAP phenomenon was in fact reproducible. Obokata was permitted to perform this closely monitored study from July 14 to November 30, 2014 under my supervision as head of the Scientific Validity Examination Team, at the direction of the Head Office for Internal Reform organized by the RIKEN President. Unfortunately, I have been unable to contact her since the completion of the study, or to obtain her agreement to be listed as an

author on this article. Nonetheless, given the extraordinary degree of attention and controversy the original STAP publications and research misconduct generated, I feel it is important to report the results of this investigation in the interests of clarifying the scientific record. In the Scientific Validity Examination Team, Hitoshi Niwa, one of the coauthors of the *Nature* papers, also conducted an independent examination of whether the STAP phenomenon was reproducible; the results of his examination have been reported previously (Niwa, 2016).

The investigation reported here consisted of two types of experiments; preliminary ones conducted without supervision, and formal ones conducted in the presence of expert witnesses. There were no significant differences in the data generated in the preliminary and formal experiments, and all are included together in this report. The experiments were conducted in a new setting, not in the laboratory that Obokata had used for the previous studies described in the retracted *Nature* publications (Obokata *et al.*, 2014a; Obokata *et al.*, 2014b). All reagents, materials, instruments, and experimental spaces were freshly furnished. Obokata was permitted to conduct experiments only in designated rooms. She prepared cell aggregates, but did not perform any of the subsequent analyses herself, other than observations of the cell aggregates by phase and fluorescence microscopy. Other members of the team conducted chimeric, FACS, qPCR and immunohistochemical analyses of the cell aggregates. In this report, I refer to the studies reported in the papers retracted (Obokata *et al.*, 2014a; Obokata *et al.*, 2014b) as "the previous studies" for the sake of brevity. I also refer to the technical tips published by several authors of the original articles for details of the experimental procedure (Obokata *et al.*, 2014c).

## Results

### Frequency of GFP-positive cells from spleen of *Oct-GFP* transgenic mice

Experiments were performed using a transgenic mouse line harboring *GFP* under an *Oct4* promoter (Ohbo *et al.*, 2003); the line is the same as that used in the previous studies (Obokata *et al.*, 2014a; Obokata *et al.*, 2014b). The mouse line has been maintained in C57BL/6 background in a homozygous state (*B6 oct-gfp<sup>+/+</sup>*). Spleens were dissected from homozygous newborn mice (*B6 oct-gfp<sup>+/+</sup>*; 6–8 days old) obtained by crossing a homozygous transgenic female (*B6 oct-gfp<sup>+/+</sup>*) with a homozygous transgenic male (*B6 oct-gfp<sup>+/+</sup>*), or from hemizygous newborn mice (*F1 oct-gfp<sup>+/-</sup>*; 6–8 days old) obtained by crossing a homozygous transgenic female (*B6 oct-gfp<sup>+/+</sup>*) with a wild type 129 male (*129 oct-gfp<sup>-/-</sup>*). Spleen cells were prepared as described previously (Obokata *et al.*, 2014a; Obokata *et al.*, 2014c), but enrichment of CD45-positive cells by FACS sorting was omitted. The source of the cells used in these experiments were lymphocytes collected with Lympholyte following the manufacturer's instructions (Cedarlane Laboratories, Ontario, Canada).

The stress treatment evaluated was the low-pH condition; no other conditions, such as trituration, were examined. The low-pH conditions included not only the previously reported induction by HCl (Obokata *et al.*, 2014a; Obokata *et al.*, 2014b; Obokata *et al.*, 2014c), but also that by ATP. Although not described in the

previous reports, the ATP treatment had been used most frequently by Obokata *et al.*, and is described in their patent application regarding the STAP process (US Patent Application no.: 14/397,080). In brief, the low-pH condition was generated by suspending the  $1 \times 10^6$  cells in 494  $\mu$ l HBSS (Hank's Balanced Salt Solution), adding 6  $\mu$ l 200 mM ATP, and incubating for 15 min at 37°C in CO<sub>2</sub> incubator. The low-pH treated-cells were cultured for 6–8 days, and cell aggregates of 50–100  $\mu$ m showing green fluorescence were identified (see Materials and methods). **Table 1** gives the frequency of the cell aggregates identified by Haruko Obokata (see Materials and methods). The frequency of green fluorescent cell aggregates was slightly higher under ATP treatment than HCl treatment in the C57BL/6 background. However, no marked difference was found in the frequency of green fluorescent cell aggregates under either of the low-pH conditions (HCl or ATP) or genetic background of mice (C57BL/6 or F1 between C57BL/6 and 129). The observed frequency was approximately 10 green fluorescent cell aggregates per  $10^6$  cells seeded; this was approximately 10-fold lower than that in the previous studies. Most green fluorescent cell aggregates also exhibited higher or lower degrees of red fluorescence (**Figure 1**). No quantitative determination was made, but about one in three cell aggregates exhibited green fluorescence more intense than red fluorescence.

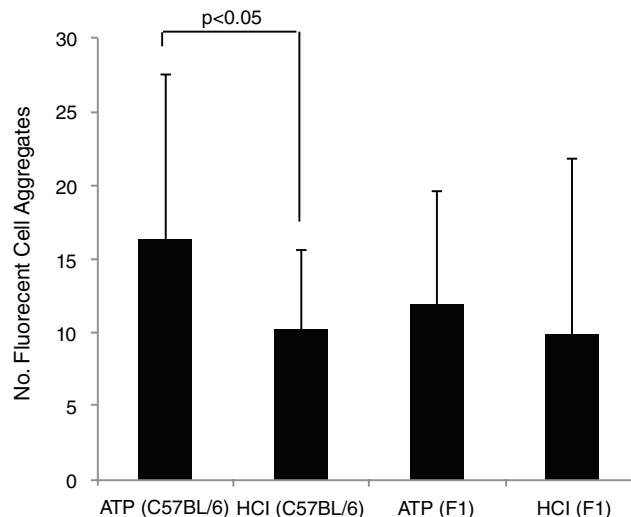
Green fluorescent cell aggregates that exhibited no significant red fluorescence were rare.

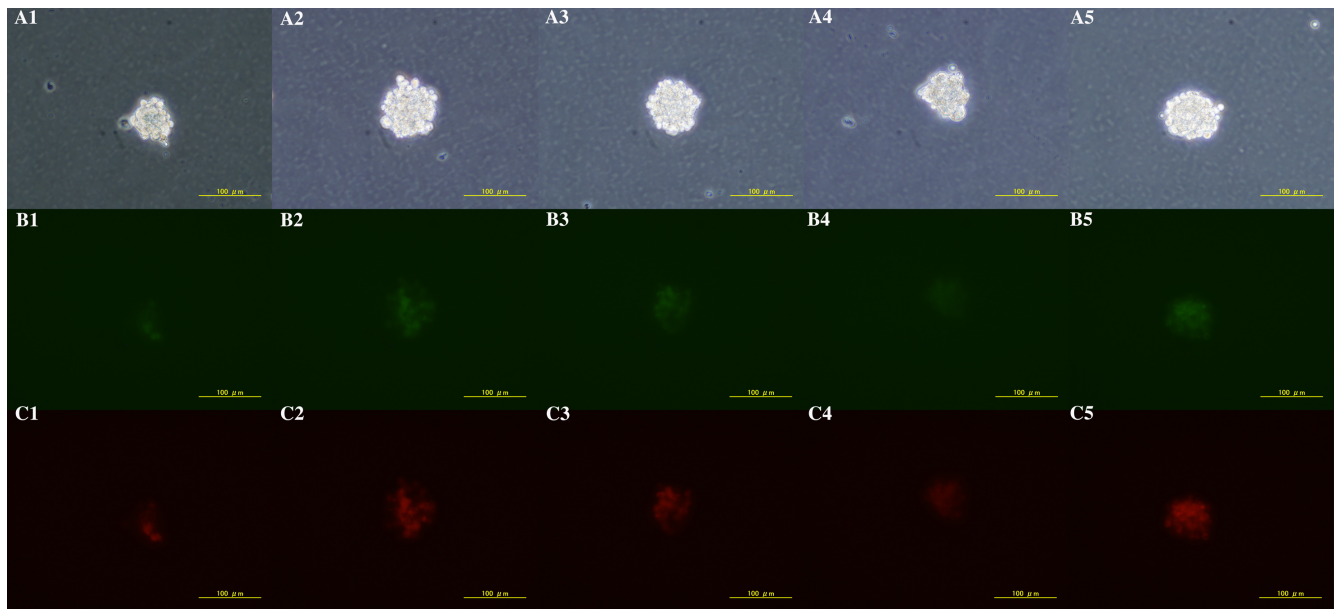
#### Chimeric potency of 'STAP' cell aggregate

Chimera production was performed with spleens of a transgenic mouse line harboring *gfp* under a *CAG* promoter (Okabe *et al.*, 1997) which were also maintained homozygously in C57BL/6 background (*B6 cag-gfp<sup>+/+</sup>*); this line is different from the one previously used (Obokata *et al.*, 2014a; Obokata *et al.*, 2014b). Spleens were dissected from *cag-gfp* hemizygous newborn mice in C57BL/6 background (*B6 cag-gfp<sup>+/-</sup>*; 6–8 days old) obtained by crosses of wild-type C57BL/6 (*B6 cag-gfp<sup>-/-</sup>*) female with homozygous transgenic (*B6 cag-gfp<sup>+/+</sup>*) male, or from hemizygous newborn mice in F1 background between C57BL/6 and 129 (*F1 cag-gfp<sup>+/-</sup>*; 6–8 days old) obtained by crosses of homozygous transgenic female (*B6 cag-gfp<sup>+/+</sup>*) with wild-type 129 male (*129 cag-gfp<sup>-/-</sup>*). Cell aggregates of 50–100  $\mu$ m were selected by their cluster morphology by Obokata and subjected to the chimeric assay. Chimeras were made by members of the Laboratory for Animal Resources and Genetic Engineering, CDB, with expertise in chimera production with ES cells (<http://www2.clst.riken.jp/arg/APDBN.html>, [http://www2.clst.riken.jp/arg/mutant\\_mice\\_generated\\_in\\_CDB.html](http://www2.clst.riken.jp/arg/mutant_mice_generated_in_CDB.html))(present affiliation: Animal Resource

**Table 1.** Frequency of cell aggregates from *oct-gfp* transgenic spleen after low pH treatment.

Mouse Background	Treatment	No. of Exp.	No. Fluorescent Cell Aggregates/ $10^6$ Cells Seeded		
			Average	Minimum	Maximum
C57BL/6	No treat.	9	0	0	0
	ATP	14	16	4	52
	HCl	11	10	3	23
F1 (C57BL/6 $\times$ 129)	No treat.	9	0	0	0
	ATP	13	12	1	30
	HCl	10	10	0	42





**Figure 1.** Examples of cell aggregates generated from *oct-gfp* transgenic spleen by low pH treatment. (A) Phase contrast views of typical five cell aggregates, (B) their green fluorescence and (C) their red fluorescence.

Development Unit, Biosystem Dynamics Group, Division of Bio-Function Dynamics Imaging, Center for Life Science Technologies (CLST)).

The previous report indicated that the generation of chimeras using STAP cells involved a distinct technical approach (Obokata *et al.*, 2014a): “Single cell dispersion by trypsinization, as it is done in the chimera production with ES cells, caused low chimaerism. STAP spherical colonies were cut into small pieces using a micro-knife under the microscope. Small clusters of the cells are then injected into blastocysts.” In the present study, cell aggregates were cut into small pieces by either glass capillary, laser beam (XY Clone: Nikko Hansen & Co., Osaka, Japan) or microsurgical knife (K-5310: FEATHER Safety Razor Co., Osaka, Japan) and were injected into a host embryo, either E2.5 8-cell stage or E3.5 blastocyst stage embryos of random-bred ICR (Charles River, Tokyo, Japan). Injected embryos were transplanted into the uterus of pseudopregnant females of the ICR strain, and recovered at E9.5 or E8.5 to judge the contribution of injected cells by GFP-green fluorescence (Table 2). Notably, the previous study describes that small clusters of ‘STAP’ cells were injected into ‘E4.5 blastocysts’, and the next day, the chimeric blastocysts were transferred into pseudopregnant females (Obokata *et al.*, 2014a).

Five hundred and sixty four embryos (210 morula and 354 blastocyst) were injected with cell aggregates cut into pieces by glass capillaries, and 294 embryos were recovered at E9.5. Ninety-two embryos (48 morula and 44 blastocyst) were injected with cell aggregates cut into pieces by laser beam, and 58 embryos were recovered at E9.5. Three hundred and ninety five embryos (193 morula and 202 blastocyst) were injected with cell aggregates cut into pieces by microknife, and 239 embryos were recovered.

Seven hundred and sixty seven embryos were injected with cell aggregates derived from C57BL/6 spleen (*B6 cag-gfp<sup>+/+</sup>*), and 284 embryos with aggregates from F1 spleen between C57BL/6 and 129 (*F1 cag-gfp<sup>+/+</sup>*). Cell aggregates cut into pieces were injected into 451 morula- and 600 blastocyst-stage embryos. In total, 1,051 embryos injected with cell aggregates cut into pieces were transplanted into a foster uterus, and 591 embryos were recovered. The contribution of injected cells was judged by GFP green fluorescence in embryos (see Materials and methods). No significant contribution of the injected cells was observed in any of the 591 embryos examined. Pluripotency was not examined by injecting putative STAP cells into tetraploid embryos.

## Discussion

Two reports of unsuccessful attempts at “STAP cell” formation have been published; both judged by *oct-gfp* and pluripotent marker gene expression (De Los Angeles *et al.*, 2015; Tang *et al.*, 2014). One of the central claims in the original reports was that the purported STAP cells had the ability to differentiate into multiple lineages, including germ cells, when placed in a normal developmental environment. However, this was not examined in these two reports. The key question is how “STAP cells” are prepared. In the original reports, the STAP cells were prepared by Haruko Obokata, while the chimera production and the establishment of two stem cells, STAP-SCs that were ES (embryonic stem)-like and FI-SCs that could differentiate into both extraembryonic and embryonic tissues, were made by Teruhiko Wakayama. Thus, Obokata is presumably the best qualified to prepare “STAP cells”. The present study focused on assessing pluripotency by chimera production using cell aggregates prepared by Obokata. Under the assay conditions reported here, we observed no evidence of pluripotency in cell aggregates she prepared herself.

**Table 2. Chimera analysis of pluripotency.**

Cell aggregates injected		No. host embryos <sup>3)</sup>		Chimera analysis		No. chimeric embryos or mice		
Genetic background <sup>1)</sup>	Specifications <sup>2)</sup>	morula	blastocyst	stage	No. developed			
C57BL/6	glass capillary	156	—	E9.5	107	0		
		—	226	E9.5	117	0		
		—	66 <sup>4)</sup>	E9.5	27	0		
F1(C57BL/6x129)		54	—	E9.5	25	0		
		—	46	E9.5	7	0		
		—	16 <sup>4)</sup>	E9.5	11	0		
C57BL/6	laser beam	30	—	E9.5	10	0		
		—	35	E9.5	26	0		
		18	—	E9.5	13	0		
F1(C57BL/6x129)		—	9	E9.5	9	0		
		C57BL/6	microknife	31	—	E8.5	23	0
				—	46	E8.5	22	0
89	—			E9.5	65	0		
F1(C57BL/6x129)		—	88	E9.5	59	0		
		47	—	E8.5	35	0		
		—	40	E8.5	17	0		
		26	—	E9.5	2	0		
		—	28	E9.5	16	0		
		Total	451	600	Total	591	0	
		1051						

1) Genetic background of *cag-gfp* mice from which spleen was isolated.

2) How cell aggregates were cut into small pieces.

3) No. embryos injected with cell aggregates and transplanted into uterus of foster mothers.

4) Embryos were transplanted into foster mothers on the next day after the injection of cell aggregates.

I encourage readers to recognize a number of limitations in the studies, which were conducted under strict time constraints and in the face of considerable, often adversarial, media scrutiny. Unfortunately, it was not possible to receive technical advice from Teruhiko Wakayama in the chimera production reported here, and it is unclear whether or to what extent the techniques for chimera production in the present study correspond to those used in the previous studies. Previous studies also examined the pluripotency of purported STAP cells by their potency to generate teratomas in immune-deficient mice. However, more than  $10^5$  cells are required to form teratoma subcutaneously in the flank of an immune-deficient mouse using ES or EC (embryo carcinoma) cells, and the process takes about one month. No teratoma formation was examined in the present study, since the frequency of green fluorescent cell aggregates was low and time was limited. Teratoma formation under the kidney capsule, which also takes about two months using blastocyst embryos, was also not examined.

The more critical question is whether and to what extent the STAP cell aggregates prepared by Obokata in this trial under new experimental conditions recapitulated the STAP cell aggregates reported in the previous study. The frequency of green fluorescent cell aggregates from low pH-treated, *oct-gfp* transgenic spleen cells was 10-fold less than that in previous studies. Moreover, green fluorescence due to GFP expression cannot be distinguished from that due to autofluorescence, nor can GFP expression by reprogramming be distinguished from that due to non-specific gene expression in dying cells. The cell aggregates were not characterized *in vitro* in detail, but the following features were observed:

(1) Preliminary FACS analysis of low pH-treated, *oct-gfp* transgenic spleen cells suggested that the frequency of green fluorescent cells was very low and that the majority of surviving cells were CD45-positive after one week in culture under the conditions used in the present study. In the previous study, CD45<sup>+</sup> cells were rare

and a significant number of green fluorescent cells were observed (Figure 1c in [Obokata et al., 2014a](#)).

(2) Preliminary qPCR analysis suggested that the majority of the cell aggregates generated in the present study did not express pluripotency markers, in contrast to the report of pluripotency marker expression in the previous study (Figure 2b in [Obokata et al., 2014a](#)), although there were cell aggregates at a low frequency that expressed one or multiple pluripotent markers including the *oct-gfp* transgene.

(3) Preliminary immunochemical analysis suggested that most of the cell aggregates in the present study did not express pluripotency markers. In contrast to the data shown in Figure 2a of the previous study, they did not express OCT4, SSEA1, NANOG and E-CADHERIN, ([Obokata et al., 2014a](#)).

The possibility cannot be excluded that the experimental conditions used in the present study in some way differed from the previously established optimum conditions for STAP induction. It is my view that it was beyond the scope of this examination to reassign each condition; a definitive answer to the question of whether the previously used conditions for inducing the STAP phenomenon can be indeed established or not must await further study. Nevertheless, I consider it is important to report that Haruko Obokata herself failed to reproduce the reported phenomenon, in that the putative STAP cells described here were unable to contribute to any tissues in a normal developmental environment.

Another reported feature of the STAP phenomenon was that while STAP cells themselves do not proliferate, two types of stem cells could be established from them: STAP-SCs and FI-SCs. However, as Obokata had no experience with these stem cell culture, she did not undertake the establishment of the secondary stem cell types in the present study.

## Materials and methods

### Animals

C57BL/6Njcl and 129X1/SvJmsSlc mice were purchased from CLEA Japan and Japan SLC, respectively. A transgenic mouse line harboring *gfp* under an *Oct4* promoter (*oct-gfp*/GOF-Tg; [Ohbo et al., 2003](#)) was provided by RIKEN BioResource Center (BRC) to CDB, and has been maintained in homozygous state under C57BL/6 background in CDB animal facility. A transgenic mouse line harboring *gfp* under a *CAG* promoter (*cag-gfp* Tg; [Okabe et al., 1997](#)) was provided to CDB by Masaru Okabe at Osaka University, and has been maintained in homozygous state under C57BL/6 background in CDB animal facility. Animals were housed in environmentally controlled rooms, and animal experiments were conducted under the institutional guidelines for Animal and Recombinant DNA Experiments that are consistent with ARRIVE guidelines. The experiments were approved by Institutional Animal Care and Use Committee of RIKEN Kobe Branch (Permit No., AH26-01).

### Preparation of cell aggregates

Newborn male mice of 6–8 days old were euthanized using carbon dioxide and then sterilized with 70% ethanol. Two spleens

were placed in a 15 ml conical tube, minced by scissors into paste, added with 5.5 ml HBSS (GIBCO 14170), mechanically dissociated using a Pasteur pipette and strained through a cell strainer (mesh size 40  $\mu$ m, FALCON 352340) into another conical tube. Five ml of Lympholyte-M (Cedarlane CL5031) was added to the bottom of the tube beneath the cell suspension, and the tube was centrifuged at 1,500 g for 20 min. The middle lymphocyte layer was transferred into another tube and centrifuged at 800 g for 10 min. The pelleted cells were suspended in 500  $\mu$ l HBSS, of which 6  $\mu$ l was subjected to the counting of cell number; in exchange 6  $\mu$ l 200 mM ATP (SIGMA 3377) or diluted HCl (10  $\mu$ l 35% HCl to 590  $\mu$ l HBSS) was added to the cell suspension.

The cell suspension was incubated at 37°C for 15 min in 5% CO<sub>2</sub> incubator, and then centrifuged at 1,500 rpm for 15 min at room temperature. After the supernatant was removed, B27 medium (DMEM/F-12 (GIBCO 11330) supplemented with 1,000 U LIF (ESGRO 1107), 2% B-27 (GIBCO 17504) and 1  $\mu$ g/ml bFGF (WAKO 060-04543) was added to the cell pellets to obtain 1 $\times$ 10<sup>6</sup> cells/ml suspension; one ml of the suspension was plated in each well of a 24 well plate (FALCON 353047) and cultured at 37°C in 5% CO<sub>2</sub> incubator for seven days to develop cell aggregates. Cell aggregates of 50–100  $\mu$ m were examined for green and red fluorescence with an Olympus Fluorescent Microscope IX51 (mirror units: Olympus U-MNIBA2 to detect green fluorescence and Olympus U-MWIG2 to detect red fluorescence), and the number of candidate aggregates were counted by Haruko Obokata. Images were taken with an Olympus DP70 camera coupled with Olympus DP Controller software (version 1.2.1.108).

### Chimeric assay for pluripotency

Cell aggregates prepared by Haruko Obokata were subjected to chimera production. Cell aggregates were cut into small pieces by either glass capillary, laser beam (XY Clone: Nikko Hansen & Co., Osaka, Japan) or microsurgical knife (K-5310: FEATHER Safety Razor Co., Osaka, Japan). The pieces were injected into host embryos of either E2.5 8-cell stage or E3.5 blastocyst stage embryos of random-bred ICR (Charles River, Tokyo, Japan). Injected embryos were transplanted into the uterus of pseudopregnant females of the ICR strain. Injection of cell aggregates and transplantation of the embryos into pseudopregnant females were performed as routinely done with ES cells (<http://www2.clst.riken.jp/arg/Methods.html>). Embryos were recovered at E9.5 or E8.5 and examined for the contribution of injected cells by detecting the presence of GFP-green fluorescence with LEICA fluorescence stereomicroscope M165FC (filter sets 10447407 and 10447408). E9.5 or E8.5 embryos of the *cag-gfp* transgenic line used for the preparation of cell aggregates served as positive control and wild type ICR embryos as negative control for the green fluorescence detection.

See [Niwa \(2016\)](#) for QPCR, immunostaining and FACS analysis.

### Data availability

**Open Science Framework: Dataset: Results of an attempt to reproduce the STAP phenomenon, doi 10.17605/OSF.IO/48f2z (Aizawa, 2016).**

### Author contributions

Shinichi Aizawa is responsible for all animal experiments. He supervised preparation of cell aggregate by Haruko Obokata and organized the chimera, FACS, qPCR and immunohistological analyses.

### Competing interests

No competing interests were disclosed.

### Grant information

This examination was supported by a grant for Scientific Validity Examination by RIKEN President's office.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### Acknowledgements

I would like to acknowledge Ms. Haruko Obokata's participation and efforts in this study. I am indebted to Dr. Hiroshi Kiyonari and Mr. Kenichi Inoue for chimera production and animal breeding, Laboratory of Animal Resources and Genetic Engineering for animal housing, Dr. Mariko Yamane for qPCR analysis and a member of Scientific Validity Examination Team for immunochemical and FACS analyses. I am also grateful to Ms. Kana Bando-Kadowaki, Dr. Go Shioi, Dr. Takaya Abe, Mr. Atsushi Katayama, Mr. Shigekazu Saitou, Mr. Akira Kimura, Mr. Naohiko Oba and Mr. Masahito Hatanaka, for their support to this examination. I deeply thank two senior witnesses outside of RIKEN and seven witnesses from BioResource Center, Center for Integrative Medical Sciences and Brain Science Institute, RIKEN. I thank Mr. Douglas Sipp for critical comments on and copyediting of this report.

### References

Aizawa S: **Dataset: Results of an attempt to reproduce the STAP phenomenon.** *Open Science Framework*. 2016.

[Data Source](#)

De Los Angeles A, Ferrari F, Fujiwara Y, *et al.*: **Failure to replicate the STAP cell phenomenon.** *Nature*. 2015; **525**(7570): E6–9.

[PubMed Abstract](#) | [Publisher Full Text](#)

Niwa H: **Investigation of the cellular reprogramming phenomenon referred to as stimulus-triggered acquisition of pluripotency (STAP).** *Scientific Reports in press. bioRxiv*. 2016.

[Publisher Full Text](#)

Obokata H, Wakayama T, Sasai Y, *et al.*: **Stimulus-triggered fate conversion of somatic cells into pluripotency.** *Nature*. 2014a; **505**(7485): 641–647. Retracted.

[PubMed Abstract](#) | [Publisher Full Text](#)

Obokata H, Sasai Y, Niwa H, *et al.*: **Bidirectional developmental potential in reprogrammed cells with acquired pluripotency.** *Nature*. 2014b; **505**(7485): 676–680. Retracted.

[PubMed Abstract](#) | [Publisher Full Text](#)

Obokata H, Sasai Y, Niwa H: **Essential technical tips for STAP cell conversion culture from somatic cells.** *Protocol Exchange*. 2014c.

[Publisher Full Text](#)

Ohbo K, Yoshida S, Ohmura M, *et al.*: **Identification and characterization of stem cells in prepubertal spermatogenesis in mice.** *Dev Biol*. 2003; **258**(1):209–225.

[PubMed Abstract](#) | [Publisher Full Text](#)

Okabe M, Ikawa M, Kominami K, *et al.*: **'Green mice' as a source of ubiquitous green cells.** *FEBS Lett*. 1997; **407**(3): 313–9.

[PubMed Abstract](#) | [Publisher Full Text](#)

Takahashi K, Yamanaka S: **Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.** *Cell*. 2006; **126**(4): 663–676.

[PubMed Abstract](#) | [Publisher Full Text](#)

Tang MK, Lo LM, Shi WT, *et al.*: **Transient acid treatment cannot induce neonatal somatic cells to become pluripotent stem cells. [version 1; referees: 2 approved].** *F1000Res*. 2014; **3**: 102.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Referee Status:  

---

## Version 2

Referee Report 17 October 2016

doi:[10.5256/f1000research.10449.r17036](https://doi.org/10.5256/f1000research.10449.r17036)



**Austin Smith**

Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, and Department of Biochemistry, University of Cambridge, Cambridge, UK

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.

---

## Version 1

Referee Report 22 August 2016

doi:[10.5256/f1000research.9395.r15460](https://doi.org/10.5256/f1000research.9395.r15460)



**Irene de Lazaro**

Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK

Two 2014 *Nature* papers authored by Obokata *et al.* described a novel protocol for the generation of pluripotent cells from somatic cells, based on their exposure to stress-triggering conditions such as mechanical disruption or acid treatment. Such process was named as stimulus-triggered acquisition of pluripotency (STAP), but the studies were soon retracted on the grounds of several confirmed experimental errors and indications of research misconduct. The present work by Dr. Aizawa describes a systematic study aiming to clarify whether STAP putative pluripotent cells – generated by Obokata under the supervision of RIKEN's Scientific Validity Examination Team – contribute to the development of the mouse embryo and hence can be considered *bona fide* pluripotent cells.

The conclusions reached in this study – that the claims for STAP cells functional pluripotency are not reproducible – are clearly supported by the data provided, since STAP cell contribution was not observed in any of the recovered embryos. The number of embryos injected with STAP cells to investigate their contribution to different tissues are satisfactorily high. In addition, the design of the study is very systematic and accounts for several potential sources of variability in the chimerism outcome (i.e. source of stress stimulus, technique utilised to cut the cell clusters before microinjection, stage of the embryos at



injection and recovery time). Nevertheless, please find below a few suggestions and clarifications that in my opinion could enrich the study by Dr Aizawa.

#### Experimental design:

- In the retracted studies by Obokata *et al.*, CD45<sup>+</sup> spleen cells were selected by FACS as the source for STAP cells generation. In the present study, sorting for CD45 was omitted, and the commercial product Lympholyte that reportedly allows specific isolation of lymphocytes was used instead. Being this an alteration of the original protocol under investigation, which may introduce differences in the nature of the starting cell population, it would be beneficial if the author could explain any reason behind this change.
- It is mentioned that the CAG-GFP transgenic mouse line utilized for chimerism experiments was different to that used by Obokata *et al.* in the previous, now retracted, studies. Was there a reason to select a different line?
- In the original Obokata *et al.* studies, clusters of STAP cells were injected into E4.5 embryos. However, in the present study the injection was performed into embryos at the E2.5 or E3.5 stage. Was this parameter changed to attempt a higher degree of chimerism? An explanation for this change in the experimental setup would also be desirable for clarity.

#### Presentation, treatment and discussion of data:

- **Frequency of cell aggregates from Oct-GFP transgenic spleen after low pH treatment (Table 1):** although it is stated in the text that “no apparent difference was found in the frequency of green fluorescence signal under either of the low-pH conditions (HCl or ATP) or genetic background of mice” this statement would be stronger if it was supported by an statistical test. Did the author conduct statistical analysis on these data?
- **Green and red fluorescence of cell aggregates:** the author seems to imply that the detected signal was a result of autofluorescence, which has been in fact pointed out by other researchers attempting to reproduce the work (Tang *et al.* 2014; De los Angeles *et al.* 2015; see last comment below). However, this is not clearly stated in the text. In addition, doubts concerning green autofluorescence could be easily vanished by using an anti-GFP antibody or by measuring the levels of GFP mRNA by qPCR or GFP protein by Western Blot. I strongly encourage the author to perform such studies if samples are still available since they would definitely clarify the issue. The inclusion in the study of cell aggregates derived from a wild-type mouse line would have also avoided this ambiguity.
- **Genetic background of mice to generate STAP cells for chimerism studies:** It is first stated in the text that the CAG-GFP mice involved in chimerism studies were bred on a C57BL/6 homozygous background. However, it is later highlighted in the text, as well as in Table 2, that both C57BL/6 and F1(C57BL/6x129) were included. Since those were also the backgrounds selected to maintain Oct-GFP transgenics for the cell aggregates assay, this is confusing needs clarification from the author.
- **Results included in Niwa, 2016:** the author refers several times to the results reported by Niwa (Niwa, 2016), who also investigated the reproducibility of the STAP phenomenon for the Scientific Validity Examination Team in RIKEN . It is however not clearly specified whether Niwa’s studies were conducted on the same STAP cells produced by Obokata under strict supervision. Such

clarification is important, since qPCR, immunostaining and FACS data are discussed, but not shown, and the reader is directed to Niwa's work.

- At least two other independent studies, carried out at institutions outside RIKEN, have aimed to clarify the STAP controversy (Tang *et al.* 2014, De los Angeles *et al.* 2015) and reached similar conclusions to those presented in this article. In particular, the issues on autofluorescence have been extensively scrutinized in De los Angeles *et al.* I believe that a brief discussion of the observations compiled in such studies would strengthen the present manuscript.
- **Typographical error:** "in the intestests of clarifying the scientific record" (page 2 of 8).

## References

1. Tang MK, Lo LM, Shi WT, Yao Y, Lee HS, Lee KK: Transient acid treatment cannot induce neonatal somatic cells to become pluripotent stem cells. *F1000Res.* 2014; **3**: 102 [PubMed Abstract](#) | [Publisher Full Text](#)
2. De Los Angeles A, Ferrari F, Fujiwara Y, Mathieu R, Lee S, Lee S, Tu H, Ross S, Chou S, Nguyen M, Wu Z, Theunissen T, Powell B, Imsoonthornruksa S, Chen J, Borkent M, Krupalnik V, Lujan E, Wernig M, Hanna J, Hochedlinger K, Pei D, Jaenisch R, Deng H, Orkin S, Park P, Daley G: Failure to replicate the STAP cell phenomenon. *Nature.* 2015; **525** (7570): E6-E9 [Publisher Full Text](#)
3. Niwa H: Investigation of the cellular reprogramming phenomenon referred to as stimulus-triggered acquisition of pluripotency (STAP). *Sci Rep.* 2016; **6**: 28003 [PubMed Abstract](#) | [Publisher Full Text](#)

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.

Author Response 24 Sep 2016

**Shinichi Aizawa**, RIKEN Center for Developmental Biology, Japan

Dear Dr. Irene de Lazaro,

I thank you for your comments. The manuscript was revised incorporating your suggestions. My responses are as follows:

1. There was no FACS cell sorter in the laboratory in which Obokata performed the set of supervised experiments reported here. She had previously obtained "STAP" cells using splenocytes prepared using Lympholyte-M, so we sought to determine whether she was able to repeat this in the present study. If she had succeeded, our plan was next to generate STAP cells using CD45+ cells sorted by FACS.
2. The origin of the *cag-gfp* transgenic mouse line used in the retracted *Nature* papers is unclear, and was not reported in the papers. Dr. Wakayama informed us that he generated the *cag-gfp* mouse line himself while at the University of Hawaii, but we did not make a formal investigation into this. The mouse line was no longer maintained in the animal facility of CDB and was not available to us. Alternatively, the *cag-gfp* mouse line may have been actually an *Acr/cag-gfp* mouse line (Nakanishi *et al.*, *Genomics* 80, 564-574 (2002)) as suggested in the report by Konno *et al* (Konno *et al.*, *Nature* 525,E4-5 (2015)). However, we

only became aware of this possibility at the time of that report, which was after the start of Obokata's replication attempt. In any case, the *cag-gfp* mouse line reportedly used in the original STAP reports is different from the *cag-gfp* mouse line (Okabe *et al.*, 1997) we used in the present study. It is nonetheless difficult to conceive how the difference in *cag-gfp* transgene might affect the efficiency of "STAP cell" production and chimera generation.

3. In Fig. 4a of the retracted *Nature* article, the embryo being injected with "STAP" cells clearly has a zona pellucida. However, E4.5 embryos typically no longer have this structure. In the absence of zona pellucida, injection is practically impossible. We note that E0 is generally defined as 0:00 am of the day when the plug is identified, and suggest that E4.5 may be a typographic error for E3.5. Alternatively, Dr. Wakayama may have artificially delayed the development of the embryo; however, this was not reported in the retracted *Nature* paper.
4. We have now included a statistical analysis (t-test), which indicates that the efficiency of cell aggregate formation is significantly different between ATP treatment and HCl treatment in the C57BL/6 background. However, the difference is slight. We have revised the manuscript accordingly (Table 1 and page 5 in the text).
5. This study focused on the multipotency of cell aggregates generated by Obokata using a chimeric assay as this was the central feature of the reported "STAP" phenomena. Given the time constraints of this study, other data were necessarily limited, as noted in the Discussion. As it was not the focus of the present study, I cannot state definitively that the red fluorescence observed was autofluorescence, although I feel that this is highly likely. RT-PCR analysis for GFP expression showed significant expression in several aggregates, but not in others that showed green fluorescence; however, these data were preliminary at best and are not presented.
6. The effects on both cell aggregate formation and chimeric potency of the spleens' genetic background were examined in the C57BL/6 and F1(C57BL6 x 129) background. It is well known that ES culture is strongly influenced by genetic background. Both of these backgrounds were used in the retracted *Nature* papers. I have now revised the manuscript (page 4 and page 6) to clarify this point.
7. The cell aggregates in Niwa's report were prepared by Niwa, not by Obokata.
8. The two reports are now cited and briefly discussed (page 8–9). These works did not examine multipotency by chimeric assay, and the most important issue of the present report is that cell aggregates prepared by Obokata herself did not exhibit multipotency in chimeric assays.

Best regards,  
Shin Aizawa

**Competing Interests:** No competing interests were disclosed.

Referee Report 13 June 2016

doi:10.5256/f1000research.9395.r14102

**Austin Smith**

Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, and Department of Biochemistry, University of Cambridge, Cambridge, UK

In this paper Dr Aizawa reports the outcome of attempts to reproduce the claim that exposure to low pH can convert splenocytes into pluripotent cells, so-called STAP cells, that are capable of colonising the mouse embryo. Although the two STAP papers have now been retracted acknowledging multiple errors and misconduct, the retraction notice does not state that the results are irreproducible but only says “we are unable to say without doubt whether the STAP-SC phenomenon is real”. This study is therefore a valuable service to the community. It is unfortunate that Ms Obokata cannot be contacted. It would be desirable if she confirmed her agreement with the findings. However, I do not think there is any requirement for her to be a co-author because she carried out the work under the explicit direction and supervision of Dr Aizawa.

The study design, results and interpretation are clearly presented. Putative STAP cell aggregates, as defined by Ms Obokata on the basis of fluorescence and/or morphology, were obtained. In a comprehensive series of micro-injections these cells were introduced into morulae or early blastocysts then transferred to recipient mice. No contribution was detected in 591 recovered embryos inspected for expression of a constitutive GFP reporter. Therefore the findings reported in the STAP papers cannot be reproduced using “STAP” cells generated by Ms Obokata in supervised conditions. This is a helpful clarification for the field.

I have a few minor suggestions and questions that could improve the clarity of the manuscript:

- In Table 1 the heading Exp No should be changed to No. of Expts and the heading No. Cell Aggregates should be No. Fluorescent Cell Aggregates.
- It is not clear from the Table or text what proportion of aggregates showed fluorescence or whether they all did. This should be stated.
- It is stated that the source of fluorescence could not be confirmed. Were no aggregates generated from wildtype splenocytes without a reporter? How intense is the green fluorescence in aggregates compared with the Oct4-GFP level in embryos or ES cells? The text should explain that red fluorescence is autofluorescence.
- For the chimaera experiments it is stated that “cell aggregates of 50-100µm were selected by their cluster morphology by Obokata”. Can “cluster morphology” be described more precisely?
- Typographical error: “cell aggregates were one cut into small pieces”.
- Were any injected embryos examined for donor cell survival/integration prior to uterine transfer?
- Could the author comment on the limit of detection (number of cells) for chimaera contribution at the stages examined using this reporter? The method “Embryos were ..... examined for the contribution of injected cells in each organ”. I assume this was in whole mount rather than dissected organs, but this should be declared.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** I was formerly Chair of the Advisory Council of RIKEN CDB.

Author Response 24 Sep 2016

**Shinichi Aizawa**, RIKEN Center for Developmental Biology, Japan

Dear Dr. Austin Smith,

I thank you for your comments. The manuscript was revised incorporating your suggestions. My responses are as follows:

1. The headings in Table 1 have been changed as suggested.
2. All *oct-gfp* cell aggregates exhibited fluorescence to some degrees.
3. No cell aggregates were generated from wild-type splenocytes. No direct comparison was made of the intensities of green fluorescence of cell aggregates with those in *oct-gfp* embryos or ES cells. I cannot state with certainty whether the green and red fluorescence was autofluorescence. RT-PCR analysis for GFP expression showed significant expression in several aggregates, but not in others that had green fluorescence; these data were very preliminary and thus are not shown. This examination focused on the multipotency of cell aggregates generated by Obokata using a chimeric assay, since this was the central feature of the STAP phenomena. Other data were only preliminary given the time constraints under which these experiments were performed, as described in Discussion.
4. To make chimeras, cell aggregates were prepared with *cag-gfp* splenocytes, thus GFP expression or green fluorescence cannot be used as a measure for the selection of cell aggregates. For this reason, they could only be selected by cell cluster morphology. In the present study, the selection was dependent entirely on Obokata's judgment. If she had succeeded, our plan was next to ask her to describe "cell cluster morphology" precisely.
5. Many embryos injected with *cag-gfp* cell aggregates at 8-cell stage and cultured for one day to the blastocyst stage were examined for the presence of green-fluorescent cells, and such cells were found to be present.
6. Chimeric extent was examined at E9.5 or E8.5 in whole mount. The retracted *Nature* papers show extensive colonization of the cells (Fig. 4 in the Article and Fig. 1 and Extended Data Fig. 1 in the Letter). The article reported eight chimeric embryos, showing more than 50% coat color contribution, of 48 chimeras obtained; these animals yielded "STAP"-derived offspring. This was the central finding in the now-retracted STAP reports. However, in the present study, no chimera equivalent to those in Fig. 4 in the article and Fig. 1 and Extended Data Fig. 1 in the Letter was obtained, nor were any chimeras obtained showing more than 50% contribution to coat pigmentation. Indeed, no chimera showing more than 1% contribution was obtained. I have now revised the text accordingly. We have not examined the limit of detection (minimum number of cells) with the *cag-gfp* mouse line used, since this was not the point of the present study. However, I believe it to be the case that if dozens of cells had been present together in any tissue, they would have been detectable in whole mount at E9.5 or E8.5.

Best regards,  
Shin Aizawa

**Competing Interests:** No competing interests were disclosed.

---