RESEARCH ARTICLE

REVISED A novel genotyping method to determine copy number in a mouse line commonly used for inducible transgene expression in brain and spinal cord [version 2; peer review: 1 approved, 2 approved with reservations]

Eleanor Hobbs*, Sarah Ryan*, Sara Rollinson, Stuart Allan , Stuart Pickering-Brown⁺

Division of Neuroscience and Experimental Psychology, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, M13 9PL, UK

* Equal contributors ⁺ Deceased author

V2 First published: 15 Oct 2020, 9:1249
<https://doi.org/10.12688/f1000research.26550.1>
 Latest published: 10 Feb 2022, 9:1249
<https://doi.org/10.12688/f1000research.26550.2>

Abstract

Background: The *NEFH-tTA* mouse has the human neurofilament heavy polypeptide promoter directing tetracycline-controlled transactivator protein (tTa) expression to the brain and spinal cord, allowing tissue-specific and doxycycline-suppressible expression of a target gene. Current genotyping protocols can only differentiate between wild-type and transgenic animals. Being able to differentiate between hemizygous and homozygous animals would be beneficial in experiment planning and reducing animal numbers.

Methods: We have identified the insertion site of the *NEFH-tTA* transgene via targeted locus amplification and next-generation sequencing. This was then used to design a multiplex PCR assay to distinguish between hemizygous and homozygous mice.

Results: The *NEFH-tTA* transgene is located on chromosome 12. Our genotyping method can identify hemizygous and homozygous mice.

Conclusions: The *NEFH-tTA* transgenic mouse line is a useful tool for studying a wide range of diseases including frontotemporal dementia and motor neuron disease, as well as other neurodevelopmental, neuromuscular or neurodegenerative disorders. We have designed and utilised a novel genotyping assay to distinguish between hemizygous and homozygous mice, involving a simple PCR assay. This is easily adaptable to a laboratory-specific protocol or machine, and will allow refinement of breeding strategies and a reduction in the number of animals that cannot be used in experiments.

Open Peer Review

Approval Status

	1	2	3
version 2 (revision) 10 Feb 2022		 view	
version 1 15 Oct 2020	 view	 view	 view

1. **Virginia Lee**, University of Pennsylvania, Philadelphia, USA

Mandana Hunter, University of Pennsylvania, Philadelphia, USA

2. **Keiichi Ishihara** , Kyoto Pharmaceutical University, Kyoto, Japan

3. **Sandy S. Pineda**, University of Sydney, Sydney, Australia

Any reports and responses or comments on the article can be found at the end of the article.

Keywords

NEFH-tTA, frontotemporal dementia, FTD, ALS, targeted locus amplification, genotyping, mouse

This article is included in the **Genomics and Genetics** gateway.

Corresponding author: Sarah Ryan (sarah.ryan@manchester.ac.uk)

Author roles: **Hobbs E:** Conceptualization, Formal Analysis, Investigation, Writing – Original Draft Preparation; **Ryan S:** Conceptualization, Investigation, Project Administration, Writing – Review & Editing; **Rollinson S:** Conceptualization, Methodology; **Allan S:** Funding Acquisition, Supervision, Writing – Review & Editing; **Pickering-Brown S:** Conceptualization, Funding Acquisition, Supervision

Competing interests: No competing interests were disclosed.

Grant information: Funding was provided by the MRC (MR/N025911/1) and MNDA (Pickering-Brown/Apr15/841-791). Both grants were awarded to SPB (PI) and SA (Co-I).

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

Copyright: © 2022 Hobbs E *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Hobbs E, Ryan S, Rollinson S *et al.* **A novel genotyping method to determine copy number in a mouse line commonly used for inducible transgene expression in brain and spinal cord [version 2; peer review: 1 approved, 2 approved with reservations]** F1000Research 2022, 9:1249 <https://doi.org/10.12688/f1000research.26550.2>

First published: 15 Oct 2020, 9:1249 <https://doi.org/10.12688/f1000research.26550.1>

REVISED Amendments from Version 1

Following constructive feedback from 3 reviewers, we have made some minor changes to this manuscript. The main changes are detailed below:

1. The phrase “copy number” has been changed to “zygosity” in the title
2. Figure 3 which was a pie chart has been removed and replaced with Table 5
3. A FASTA file showing the sequence of the transgene has been added to the Underlying data repository, which can be found here <https://doi.org/10.6084/m9.figshare.16847170.v1>

Any further responses from the reviewers can be found at the end of the article

Abbreviations

NEFH, neurofilament heavy polypeptide promoter; TG, transgene; tTA, tetracycline-controlled transactivator protein; TRE, tetracycline response element; TLA, targeted locus amplification.

Introduction

The *NEFH-tTA* mouse has the human neurofilament heavy polypeptide promoter directing tetracycline-controlled transactivator protein (tTA) expression to large-calibre axons of the brain and spinal cord, allowing tissue-specific and doxycycline-suppressible expression of a target gene¹. The line was generated via pronuclear injection of a plasmid with the human *NEFH* promoter isolated from BAC clone (CHORI: RP11-91J21) to drive expression of the tetracycline transactivator gene, which randomly integrated into the genome¹. While a genotyping protocol is available to distinguish between wild-type and transgenic animals¹, ideally a colony of homozygous animals would be maintained for experimental breeding with another transgenic line with the gene of interest under the control of a tetracycline response element (TRE). All offspring of these breeds would therefore be *NEFH-tTA*^{+/−} (which has been shown to be a sufficient driver of expression¹); this would allow use of littermate controls and reduce required breeding numbers as encouraged by the ARRIVE guidelines². In order to address this, we have identified the insertion site of the *NEFH-tTA*

transgene, designed a novel PCR genotyping strategy, and demonstrated its reliability in a colony of *NEFH-tTA* mice.

Methods**Ethical statement**

All animal work was approved following local ethical review by the University of Manchester Animal Welfare and Ethical Review Board and performed under Home Office project license 70/8903 and in accordance with the Home Office (Animals) Scientific Procedures Act (1986). All efforts were made to ameliorate harm to animals; mice were housed and maintained according to standard practices in the University of Manchester Biological Services Facility (detailed below) and no licenced procedures were performed.

Maintenance of transgenic mice

B6;C3-Tg(*NEFH-tTA*)8Vle/J mice (JAX stock #025397) were obtained from Jackson Laboratories (Bar Harbor, ME, USA) and bred with C57BL6/J mice to produce a colony of hemizygous *NEFH-tTA* animals (13 transgenic mice were crossed with 12 wild-type in total, as this number was sufficient to create a new colony of backcrossed mice). All animal work was carried out in the University of Manchester Biological Services Facility. Mice had free access to food and water and were housed under light-humidity- and temperature-controlled conditions: ambient temperatures of 21°C (±2°C), humidity of 40–50%, 12h light cycle, ad libitum access to water and standard rodent chow. Animals were housed up to five per cage with environmental enrichment. For breeding, mice were housed as either breeding pairs or trios of two female and one male, and pups housed with parents until weaning (approximately four weeks post-birth).

At weaning, animals were restrained by scruffing and ear biopsies were taken using a standard ear puncher and genotyped as described¹. Briefly, tissue was disrupted in proteinase K for one hour at 57°C; reactions containing 3μL of diluted DNA, 10μL PrimeSTAR HS Premix (Takara, Clontech #R040A) and 0.5μM each of forward and reverse transgene primers and internal positive control primers (Primers 1–4, Table 1) in a total volume of 20μL were amplified in a thermal cycler (ThermoFisher SimpliAmp™ A24812; Table 2). PCR products were visualised on a 2% agarose gel with HyperLadder IV (Bioline) to determine product size.

Table 1. Genotyping primer sequences.

No.	Sequence	
1	CTC GCG CAC CTG CTG AAT	Transgene allele detection ¹
2	CAG TAC AGG GTA GGC TGC TC	Transgene allele detection
3	CTA GGC CAC AGA ATT GAA AGA TCT	Internal control ¹
4	GTA GGT GGA AAT TCT AGC ATC ATC C	Internal control
5	TCATGGAGACCAATGACG	Wild-type detection
6	GGAGGAATTCATGTGCCACT	Wild-type detection

Identification of the transgene insertion point

A combination of targeted locus amplification (TLA) and next-generation sequencing was used to identify the insertion point of the *NEFH-tTA* transgene (the FASTA of the transgene sequence can be viewed in the Underlying data). This was performed by Cergentis B.V (Utrecht, The Netherlands). One mouse spleen was dissected from a six-week old *NEFH-tTA*^{+/+} mouse (culled by Schedule 1 CO₂ inhalation) and cells isolated for TLA sample prep. Spleen tissue was disrupted by gently pushing through a 40µm mesh and cells collected in PBS with 10% foetal calf serum (Gibco #10500064). Cells were pelleted by centrifugation at 250g for ten minutes at room temperature, resuspended in lysis buffer (0.15M NH₄CL, 0.01M KHCO₃, 0.0002M EDTA) and incubated at room temperature for five minutes before additional centrifugation at 250g for ten minutes. Cells were resuspended in PBS with 10% foetal calf serum (Gibco #10500064), stored at -80°C and shipped to Cergentis on dry ice.

Two primer sets (Table 3) were used in individual TLA amplifications. PCR products were purified and library prepped using the Illumina NexteraXT protocol and sequenced on an Illumina sequencer. Reads were mapped using BWA-SW, which is a Smith-Waterman alignment tool. This allows partial mapping, which is optimally suited for identifying break-spanning reads. The mouse genome version mm10 (GenBank assembly accession: [GCA_000001635.2](#)) was used for mapping.

Table 2. PCR reaction conditions.

Stage	Temperature	Time	
1	95°C	20 sec	
2 10x	95°C	20 sec	
	65°C	15 sec	- 0.5°C per cycle
	68°C	10 sec	
3 28x	95°C	15 sec	
	62°C	15 sec	
	70°C	30 sec	
4	72°C	2 minutes	
	4°C	∞	

Genotyping of offspring from hemizygous intercross matings

Hemizygous mice were bred together (total 25 animals; the number of breeding animals was chosen based on expected number of homozygous offspring and numbers required to set up a large colony to breed with a different transgenic line³) and ear biopsies taken from all offspring at weaning (n=50) as described above. Primers were designed using [Primer3](#) (web version 4.1.0) to produce a product spanning the insertion site on chromosome 12 (Primers 5-6, Table 1) and PCR reactions were set up as described under *Maintenance of transgenic mice*. REDExtract-N-Amp™ 2X PCR Reaction Mix (Sigma-Aldrich, #R4775) was used instead of PrimeStar.

Statistical analysis

Outcome vs Expected for genotype was performed using GraphPad Prism version 8.00 for Mac (GraphPad Software, La Jolla California USA, www.graphpad.com).

Results

The NEFH-tTA transgene is located on chromosome 12

Using TLA highest coverage is observed on the sequences directly surrounding the location of the primer set. Here, high coverage is observed on chromosome 12, outlined in red in Figure 1A, indicating the transgene (TG) has integrated in chromosome 12⁴.

From locus-wide coverage (Figure 1B) we were able to identify the junction sites between the TG and the mouse genomic sequence, and it was concluded that the TG has integrated in mouse chr12:6917896-chr12:6917912. Reads marking the TG integration are identified in Table 4. The 11 bases between chr12:6917896-chr12:6917912 have been deleted following the integration event. According to the reference sequence (mm10) there are no genes annotated in this region. Complete coverage is observed across the whole TG sequence from TG:68-18543 (Figure 1C). At the edges of the coverage, fusion reads have been found, connecting these ends, but the TG: 1 -68 and TG 18544 – end regions are not integrated into the mouse genome (Table 4). These regions contained sequences for bacterial DNA replication and therefore are not required for transgenic expression in the mouse.

Next to the head-to-tail fusion as reported above, four other TG-TG fusions were found within the TG sequence (Table 4). While an exact copy number cannot be determined using TLA, an estimation can be made based on the number of integration sites, number of fusion reads and the ratio of coverage on

Table 3. Targeted locus amplification primers.

Primer Set	Direction	Binding pos.	Sequence
1	Fw	6590	GACAGAATAGCAGTACCCAC
	Rv	6428	GACAAGGTTCTAGCAGTATTAAT
2	Fw	16557	TCTCTGCTCTCCGTTCTATGGA
	Rv	16206	GCTGGGCTGAGTTTGAGGG

Figure 1. (A) Targeted locus amplification (TLA) sequence coverage across the mouse genome using primer set 1. The chromosomes are indicated on the y-axis, the chromosomal position on the x-axis. Similar results were obtained with primer set 2 (see *Underlying data*⁴). (B) TLA sequence coverage across mouse chromosome 12 with primer pair 1. The arrow points towards the integration site. (C) TLA sequence coverage across the TG with primer pair 1 (top) and 2 (bottom).

Table 4. Transgenic insertion sites and fusion sequences. Homologous sequences are shown in purple.

	Join site 1	Join site 2	Sequence
5' integration	chr12:6917896 (-)	TG:15950 (-)	CCTTGGCAGGGTTGTAGATGAATGTCATCATGGAGACACCAATGACGGGCACAATGGCAGTCAA ATGACCATCACACCAAGGGAT GTCCACTTTGAGGGGTCTCTGGCTGAACTGGGATGGGAA CCACATCTTGCCGTTCTGTCAA
3' integration	TG:7482 (-)	chr12:6917912 (+)	CGCTCTTCAATCTTCCGAGGCCGATGAAGCATTGCGGGCTTCCCACTGCGGAAGGGCGGGG ATGGCTGTGACGAGGGCGTGCCCGCCGTGCTGCTAACTTATCATCACTTGAATTGATGGCAT AGCTTCTGTTGGTGAACCAATGACAGTTTTTAATCTATTTCTGGGATCTTCTCCAAAATTGCAA AAGATAGTTTAAATCATATGTGTAACATG
Fusion read	TG:18537 (-)	TG:74 (+)	CCTCATCTCCATGTCCTTCTCTGTGGAGCCGTGTGCGCTGCACATATCTAAGTACATGCTGG GGATGCGGTTGGCTCTATGGCTTCTGAGGCGGAAAGAACCACTGGGGCTCGA ACGGCTTGACATTGATTGTGACTAGTTATTAAAT
TG-TG fusion 1	TG:7926 (-)	TG:70 (+)	GATTAGAGGTGTGAGCCACTGCACCCGCGAAAGTGGTTTGATGACAATGTTGTGAGGCTAGCGTTGACATTGATTATTAGTTATTAAATAG
TG-TG fusion 2	TG:3745 (+)	TG:4793 (+)	ATCTCAACTCACTGCAGCCTTGACCTCTCTGGCTCAAGCGATCCACCCACTTCAT CTCTCCTAGTAGTTGGGACTACAGGTGCCTGCCAACACACC
TG-TG fusion 3	TG:7576 (-)	TG:7714 (-)	GGGGGCGCTGTTGCTGGCGCGATTCTATCAAGCGCAGAGCCAGCCGAGGGCAGTATGCCGAGG CAGAGAAACCGGTTCTGTCTCCAGGG CTCGGG GAGGCTGAGGTGGGAGGATCGCTTAAAC CTGGAGAGGCGGAGGTTGCAGTGAGCCGAGATCACGCCACGGCACTCCAGCTCGGCAACAG AGCGAGACCCTGTACAAACGAACA
TG-TG fusion 4	TG:11034 (-)	TG:11295 (-)	TTTAACTCTGGCATGAATCTTCTCTCAAAATGTACTATGGGCTCCCTCTCTCTCCCCAGAATA AAGAGATG AAACAGC TTTGTATCTCTTACGTGGTGGGCACCTGCTGCTGTGCTCCACGTACATACGTGGTGTCTGCAACAACCTGTGCAGCTGTGCAA GGTCACTGTCCTCTAC

the TG and genome integration site. Here, one integration site was found and a total of five TG-TG fusions were found. The coverage at the TG was significantly higher compared to the genome. These facts together suggest that >5 copies of the TG may have integrated.

Design of a PCR assay to distinguish between hemizygous and homozygous transgenic mice

We designed primers to amplify a wild-type PCR product of 488bp from chromosome 12, which would be disrupted by insertion of the transgenic allele (Figure 2A). A multiplex PCR

reaction including the original primers to the transgene¹ (Table 1) was carried out on samples from a hemizygous intercross and proved able to distinguish between wild-type (488bp band only), hemizygous (both bands) and homozygous (150bp band only) (Figure 2B)⁴.

Homozygous NEFH-tTA mice are observed at the same rate as expected

After genotyping 50 offspring of hemizygous intercrosses we observed no significant difference in the number of *Nefh-tTA*^{+/+} mice compared to the expected ratio (Table 5,

Figure 2. Genotyping strategy for homozygous mice. (A) Primer design to produce a wild-type product that is disrupted in the presence of the transgenic allele. (B) Samples visualised on a 2% agarose gel showing wild-type (top band), homozygous (bottom band) and hemizygous (both bands) animals.

Table 5. Numbers of pups of each genotype observed using our novel PCR assay in litters from hemizygous intercrosses compared to numbers which would be expected based on Mendelian inheritance patterns. A chi-square test was performed (using GraphPad Prism version 7.04) to confirm that the observed numbers were not significantly different to expected, indicating that our genotyping method works as predicted.

Genotype	Number expected	Number observed	Percentage of total expected	Percentage of total observed
<i>Nefh-tTA</i> ^{-/-}	12.5	16	25	32
<i>Nefh-tTA</i> ^{+/-}	25	26	50	52
<i>Nefh-tTA</i> ^{+/+}	12.5	8	25	16
TOTAL:	50	50	100	100
Chi-square value:	2.64			
P value (two-tailed):	0.2671			
Significance:	Not significant			

chi-square=2.64, $p < 0.2671^4$), confirming our novel genotyping assay performs as expected. Homozygous mice grew to adulthood and did not exhibit any overt phenotype.

Discussion/conclusion

The *NEFH-tTA* transgenic mouse line is a useful tool for studying a wide range of diseases including frontotemporal dementia and motor neurone disease, as well as other neurodevelopmental, neuromuscular or neurodegenerative disorders. Here we have shown that the transgene has inserted into chromosome 12. Furthermore, we have designed and utilised a novel genotyping assay to distinguish between hemizygous and homozygous mice, involving a simple PCR assay. This is easily adaptable to a laboratory-specific protocol or machine and will allow researchers using this line to refine their breeding strategies and reduce the number of animals that cannot be used in experiments.

Data availability

Underlying data

Figshare: "A novel genotyping method to determine zygosity in a mouse line commonly used for inducible transgene expression in brain and spinal cord – Raw data/analysis" DOI: <https://doi.org/10.6084/m9.figshare.12982220.v1>⁴.

This project contains the following underlying data:

- Nefh o vs e.pzfx (raw data and analysis for Table 5)

- Nefh o vs e.xlsx (raw data and analysis for Table 5)
- TLA and seq data – Cergentis report 2017UMAN0801.pdf (all sequencing and TLA data received from Cergentis)
- Raw data – PCR gel image.tif (original, unedited PCR gel image from Figure 2)

Figshare: FASTA reference file for the NEFH-tTA transgene, <https://doi.org/10.6084/m9.figshare.16847170.v1>⁵

This project contains the following underlying data:

- FASTA reference file for the NEFH.docx

Data are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Acknowledgements

We would like to thank the University of Manchester Biological Services Facility for animal care.

Prof Stuart Pickering-Brown, the last author listed on this article, has sadly passed away since submission of the original version. This publication is in his memory.

References

1. Walker AK, Spiller KJ, Ge G, *et al.*: **Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43.** *Acta Neuropathol.* 2015; **130**(5): 643–660.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Kilkenny C, Browne WJ, Cuthill IC, *et al.*: **Improving bioscience research reporting: The arrive guidelines for reporting animal research.** *Animals.* 2014; **4**(1): 35–44.
[Publisher Full Text](#) | [Free Full Text](#)
3. Ryan S, Hobbs E, Rollinson S, *et al.*: **CRISPR/Cas9 does not facilitate stable expression of long C9orf72 dipeptides in mice.** *Neurobiol Aging.* 2019; **84**: 235.e1–235.e8.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Ryan S, Hobbs E, Pickering-Brown S, *et al.*: **Raw data/analysis.** *figshare.* Dataset, 2020.
<http://www.doi.org/10.6084/m9.figshare.12982220.v1>
5. Ryan S: **FASTA reference file for the NEFH-tTA transgene.** *figshare.* Dataset, 2021.
<http://www.doi.org/10.6084/m9.figshare.16847170.v1>

Open Peer Review

Current Peer Review Status: ? ✓ ?

Version 2

Reviewer Report 31 March 2022

<https://doi.org/10.5256/f1000research.116746.r123131>

© 2022 Ishihara K. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Keiichi Ishihara

Department of Pathological Biochemistry, Division of Pathological Sciences, Kyoto Pharmaceutical University, Kyoto, Japan

The manuscript has been improved.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neuroscience, Pathological biochemistry

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

Version 1

Reviewer Report 28 April 2021

<https://doi.org/10.5256/f1000research.29310.r79860>

© 2021 Pineda S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sandy S. Pineda

¹ Brain and Mind Centre, University of Sydney, Sydney, Australia

² Brain and Mind Centre, University of Sydney, Sydney, Australia

In the manuscript entitled: "A novel genotyping method to determine copy number in a mouse line commonly used for inducible transgene expression in brain and spinal cord", by Hobbs and

colleagues, they present the use of TLA (Targeted locus amplification) analysis to identify the integration site of the NEFH-tTA transgene in B6;C3-Tg(NEFHTTA)8Vle/J mice, and develop a simple genotyping assay to distinguish Heterozygous from Homozygous animals for the transgene. The use of TLA allowed the mapping of the random integration site to ch12qA1.1 in the mouse genome. The study is quite simple, but clear in their objective of providing researchers a new genotyping protocol that would facilitate the handling of the mouse colony and planning of future experiments.

However, there is some minor changes that would improve the manuscript:

- Could the authors include an extended materials and methods sections to include a summary of the TLA method and the bioinformatics commands used? A summary figure, could also be included to summarise the method.
- Since TLA was used to get the exact location of the integration, a supplementary FASTA file with the entire sequence including the junctions might help the reader to better understand the insertion site.
- I would have suggested to do a high-resolution copy number analysis, as an extra step. I find the use of copy number in the title somewhat misleading, since the only confirmation of the copy number comes indirectly from the TLA analysis.
- I do not see the reason for figure 3, it would be best to be removed or to be merged to figure 2. If kept, you will also need to provide more details from the statistical analysis done, maybe adding the table right next to it.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neuroscience.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 17 Dec 2021

Sarah Ryan

We thank Dr Pineda for reviewing the manuscript, and are grateful for the constructive comments. There were some points to be addressed, as follows:

Could the authors include an extended materials and methods sections to include a summary of the TLA method and the bioinformatics commands used? A summary figure, could also be included to summarise the method.

Unfortunately this information is proprietary as a company (Cergentis) performed this analysis for us. However, we have included the full report with all the information provided by Cergentis as part of the Underlying Data repository.

Since TLA was used to get the exact location of the integration, a supplementary FASTA file with the entire sequence including the junctions might help the reader to better understand the insertion site.

We thank the reviewer for this useful comment and agree the sequence should be included. Since the FASTA is a long text file we have uploaded it as part of the extended/underlying data repository accompanying the article and referred to it in the main text.

I would have suggested to do a high-resolution copy number analysis, as an extra step. I find the use of copy number in the title somewhat misleading, since the only confirmation of the copy number comes indirectly from the TLA analysis.

We appreciate that our statements regarding copy number were too definitive, and have changed the title to reflect this. We have also removed mention of multiple copies from the conclusion and edited the results section to clarify that multiple copies are a possibility but that we cannot definitively conclude this from this study.

I do not see the reason for figure 3, it would be best to be removed or to be merged to figure 2. If kept, you will also need to provide more details from the statistical analysis done, maybe adding the table right next to it.

We have removed Figure 3 and replaced with Table 5, as suggested. The GraphPad Prism file containing the full analysis is also available in the Underlying Data repository.

Competing Interests: NA

Reviewer Report 28 April 2021

<https://doi.org/10.5256/f1000research.29310.r83325>

© 2021 Ishihara K. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Keiichi Ishihara

¹ Department of Pathological Biochemistry, Division of Pathological Sciences, Kyoto Pharmaceutical University, Kyoto, Japan

² Department of Pathological Biochemistry, Division of Pathological Sciences, Kyoto Pharmaceutical University, Kyoto, Japan

Eleanor Hobbs *et al.* in this paper, have identified the integrated site of the transgene in the B6;C3-Tg(NEFH-tTA)8Vle/J mice by TLA and also have developed a simple PCR method to distinguish for hemizygous and homozygous. The manuscript is well written, but some typos, infelicities and concerns remain in existence as following:

Abbreviation: "tTa" correct to "tTA".

Introduction: Definition of "tTA" is duplicated.

Methods: How many times were the transgenic mice backcrossed to the C57BL/6J mice before they were used in this study?

Results:

- In the first paragraph, transgene is abbreviated as TG, but the word, transgene is already appeared in the introduction.
- The figure 3 is better to described as a table, showing number and ratio in each genotype.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neuroscience, Pathological biochemistry

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 17 Dec 2021

Sarah Ryan

We thank Prof Ishihara for reviewing the manuscript, and are grateful for the constructive comments. There were some points to be addressed, as follows:

How many times were the transgenic mice backcrossed to the C57BL/6J mice before they were used in this study?

Our standard breeding strategy for maintenance of the transgenic line before the genotyping assay was developed was to always cross transgenic mice with wild-type. This was because the pre-existing genotyping assay could only distinguish between transgenic and wild-type but not homo/hemizygous transgenics, and it had previously been reported that homozygous transgenics of this line do not breed well together. Therefore all experimental mice were backcrossed to C57BL/6J several times before use in this study. However, we did not have a specific backcrossing strategy since the transgene is an expression system rather than a particular gene of interest, and we were not investigating any phenotypes in these mice.

The figure 3 is better to described as a table, showing number and ratio in each genotype

We have removed Figure 3 and replaced with Table 5, as suggested. The GraphPad Prism file containing the full analysis is also available in the Underlying Data repository.

Three typing errors were also noted, which we have corrected.

Competing Interests: NA

Reviewer Report 29 October 2020

<https://doi.org/10.5256/f1000research.29310.r73226>

© 2020 Lee V et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Virginia Lee

¹ Department of Pathology and Laboratory Medicine, Institute on Aging and Center for

Neurodegenerative Disease Research, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

² Department of Pathology and Laboratory Medicine, Institute on Aging and Center for Neurodegenerative Disease Research, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Mandana Hunter

¹ University of Pennsylvania, Philadelphia, PA, USA

² University of Pennsylvania, Philadelphia, PA, USA

In this manuscript from Eleanor Hobbs et al., the authors report the use of targeted locus amplification to identify the integration site of the NEFH-tTA transgene in B6;C3-Tg(NEFH-TTA)8Vle/J mice, which have been used in disease models for neuron-specific expression of doxycycline-regulatable cassettes in the brain and spinal cord. Having mapped the NEFH-tTA integration to ch12qA1.1, the authors proceed to develop a simple PCR assay to discriminate mice hemizygous and homozygous for the transgene.

The study is well performed and lucidly communicated. The reported PCR assay is anticipated to be of use to investigators breeding B6;C3-Tg(NEFH-TTA)8Vle/J mice, facilitating crossing to generate homozygotes. Some minor amendments would enhance the manuscript and these are discussed below:

- The reference to 'copy number' in the title is misleading since the reported genotyping method actually assesses zygosity of the NEFH-tTA locus rather than NEFH-tTA transgene copy number per se.
- The amplicon reads from targeted locus amplification are aligned to GRCm38 despite an updated assembly - GRCm39 – now being available. While it is unnecessary to repeat the bioinformatic analyses, it would be useful to add the NEFH-tTA 5' and 3' integration coordinates in terms of the latest *Mm* genome build.
- It would be helpful to provide a FASTA reference file for the NEFH-tTA sequence so that readers can more easily interpret the junction and fusion sites identified in Table 4.
- The manuscript indicates that similar targeted locus amplification findings were found for the second primer set and references the online 'Underlying data' files (Fig. 1A legend). However, data showing the alignment of NGS reads by chromosome for primer set 2 are not actually in the supplemental files. This would be best resolved by adding the missing data or, less optimally, by removing reference to the second primer pair or clarifying what is meant by "similar results".
- The molecular biology aspects of the targeted locus amplification method are not described. Moreover, the Illumina instrument used should be identified and the 'bwa bwasw' alignment parameters should be given.
- The second paragraph of the Results section indicates that the NEFH-tTA integration coordinates were identified from locus-wide coverage data (Fig. 1B), but presumably, it was actually the sequence of the ch12/TG junction reads that identified the integration site.

- The treatment of data pertaining to the NEFH-tTA integration copy number in the ch12 locus is not entirely clear (results paragraphs 2 and 3). Is it possible that any of the TG-TG fusion amplicons were generated secondary to the fragmentation and re-ligation procedure used in targeted locus amplification? Without orthogonal sequencing of the locus, the reader is left unclear as to the proposed arrangement of the putative multiple transgene copies at the integration site. Left unresolved, the corresponding statement in the conclusion (“...we have shown that multiple copies of the transgene have inserted...”) may be viewed as excessively definitive.

Figure 3 is superfluous and could be replaced by the Chi-square dependency table.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 17 Dec 2021

Sarah Ryan

We thank Prof Lee for reviewing the manuscript, and are grateful for the constructive comments. There were some points to be addressed, as follows:

The reference to ‘copy number’ in the title is misleading since the reported genotyping method actually assesses zygosity of the NEFH-tTA locus rather than NEFH-tTA transgene copy number per se.

We agree this was incorrect use of the phrase “copy number” and have changed the title accordingly.

The amplicon reads from targeted locus amplification are aligned to GRCm38 despite an updated assembly - GRCm39 – now being available. While it is unnecessary to repeat the bioinformatic analyses, it would be useful to add the NEFH-tTA 5' and 3' integration coordinates in terms of the latest Mm genome build.

A BLAST of the integration site sequences using the updated GRCm39 demonstrates that these coordinates are the same as those listed in the manuscript, and therefore no update is necessary in this case.

It would be helpful to provide a FASTA reference file for the NEFH-tTA sequence so that readers can more easily interpret the junction and fusion sites identified in Table 4.

We thank the reviewer for this useful suggestion and agree this should be included. Since the FASTA is a long text file we have uploaded it as part of the extended/underlying data repository accompanying the article and referred to it in the main text.

The manuscript indicates that similar targeted locus amplification findings were found for the second primer set and references the online 'Underlying data' files (Fig. 1A legend). However, data showing the alignment of NGS reads by chromosome for primer set 2 are not actually in the supplemental files. This would be best resolved by adding the missing data or, less optimally, by removing reference to the second primer pair or clarifying what is meant by "similar results".

The reason this additional data is not in supplemental files is because F1000 does not allow supplementary data to be added to articles, but rather makes use of the "underlying data" repository to include both raw data and extended data. Therefore the reference to underlying data in this statement should be considered as equivalent to a reference to a supplemental figure in a different journal.

The molecular biology aspects of the targeted locus amplification method are not described. Moreover, the Illumina instrument used should be identified and the 'bwa bwasw' alignment parameters should be given.

Unfortunately we are unable to provide this information as Cergentis performed this part of the analysis using their proprietary methods. However, we have included the full report provided by Cergentis as part of the Underlying Data.

The second paragraph of the Results section indicates that the NEFH-tTA integration coordinates were identified from locus-wide coverage data (Fig. 1B), but presumably, it was actually the sequence of the ch12/TG junction reads that identified the integration site.

We have edited the text here to further clarify that the locus-wide coverage data allowed us to identify the junction sites and this in turn identified the integration site on chromosome 12.

The treatment of data pertaining to the NEFH-tTA integration copy number in the ch12 locus is not entirely clear (results paragraphs 2 and 3). Is it possible that any of the TG-TG fusion amplicons were generated secondary to the fragmentation and re-ligation procedure used in targeted locus amplification? Without orthogonal sequencing of the locus, the reader is left unclear as to the proposed arrangement of the putative multiple transgene copies at the integration site. Left unresolved, the corresponding statement in the conclusion ("...we have shown that multiple copies of the transgene have inserted...")

may be viewed as excessively definitive

We have removed this statement from the conclusion, and edited a similar comment in the results section to clarify this is a possibility but not definitive.

Figure 3 is superfluous and could be replaced by the Chi-square dependency table.

We have removed Figure 3 and replaced with Table 5, as suggested. The GraphPad Prism file containing the full analysis is also available in the Underlying Data repository.

Competing Interests: NA

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research