THE EMBO JOURNAL

Manuscript EMBO-2017-96603

Loss of pollen-specific phospholipase Not Like Dad (NLD) triggers gynogenesis in maize

Laurine Gilles, Abdelsabour Khaled, Jean-Baptiste Laffaire, Sandrine Chaignon, Ghislaine Gendrot, Jérôme Laplaige, Hélène Bergès, Gensèric Beydon, Vincent Bayle, Pierre Barret, Jordi Comadran, Jean-Pierre Martinant, Peter M. Rogowsky and Thomas Widiez

Corresponding author: Thomas Widiez, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA

Review timeline:

Submission date:	31 January 2017
Editorial Decision:	08 February 2017
Revision received:	08 February 2017
Accepted:	09 February 2017

Editor: Andrea Leibfried

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 08 February 2017

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by two referees whose comments are enclosed. As you will see, both referees express interest in your manuscript and are broadly in favour of publication, pending satisfactory minor revision.

I would thus like to invite you to provide a point-by-point response to the comments made by the two referees and to address them in your manuscript. Please also make sure to mention all aliases for NLD.

REFEREE REPORTS

Referee #1:

This manuscript describes the mechanism of gynogenesis, in which a haploid embryo is formed from an unfertilized egg cell, upon receiving a trigger from the sperm. In maize, this has been widely used in the production of inbred lines for breeding, critical for the success of heterosis. This work has mapped a QTL to a gene named NLD, which encodes a phospholipase. Maize inducer lines carry a mutation in the last exon of this gene, and can be complemented by a wild type construct. GUS and fluorescent protein fusions are used to show that NLD is expressed only in mature pollen and in the pollen tube, where is localizes to the sperm cell plasma membrane.

This is a particularly exciting discovery, because it could be used to induce haploids, and consequently inbred formation in other crops, as well as to improve the process in maize. Although the mechanism is completely obscure, the significance and historical value warrants publication.

Overall the paper is very well written, with just a few typos.

It is somewhat unfortunate that the QTL already has two different names, but now with the cloning is gaining a third name-and yet a fourth and fifth name with the recently published Nature and Molecular Plant papers.

In figure 1, the schematic representations of the mutant protein is misleading, the domain before and after the patatin domain is mostly identical, but shown as a different color, also the gray and dark green boxes are not so obvious to me

Figure 3, the plasma membrane localization is not clear, and could be verified by plasmolysis. Also it should be backed up by an analysis of the protein sequence, does it have a signal peptide? Is there a C-terminal membrane anchoring motif?

Page 4, high homology does not make sense, homology is absolute, not a relative term

Page 5, the word densification should be translated into standard English

Referee #2:

This manuscript describes the cloning of the maize NLD gene, together with very preliminary functional characterization of the gene and protein. The identification by fine mapping and transgenic complementation is convincing, although I would have liked to see complementation with the genomic construct rather than overexpression. I am not a cell biology expert, but the protein localization data in Arabidopsis and maize look reasonable as well, with the caveat that functionality of the construct through complementation of the mutant was not confirmed.

The data in this manuscript are at a similar level as the recent Nature paper by Kelliher et al. While that paper included addition RNA-seq data, they were not very informative and did not add much to that paper.

Overall, both papers are of a level that one would expect in Genetics, G3, Heredity, New Phytologist, given that they do not do much more than report cloning of a mutant gene. Compared to the standard in the field, especially the raft of rice QTL cloning papers over the past few years, many of them from Chinese colleagues, papers that typically had very extensive and impressive functional data, neither the Nature paper nor the EMBO manuscript can compete. I assume Nature published the other work, despite its preliminary nature, because of potential biotechnological application. While these papers are not the first to claim that they provide a new inroad to haploid induction, the simple genetic architecture of the system makes it not unlikely that this will indeed be more widely applicable than previous technologies, with the caveat that genetic background might influence the efficacy of haploid induction.

The current manuscript is obviously very rushed, but I suspect that the authors, to their credit, were holding out for more functional data. It seems that the authors felt merely compelled by the weak Nature paper to push this out.

Minor comments: The details on fine mapping etc. do not belong in the main text.

Response to the reviewers:

Referee #1:

This manuscript describes the mechanism of gynogenesis, in which a haploid embryo is formed from an unfertilized egg cell, upon receiving a trigger from the sperm. In maize, this has been widely used in the production of inbred lines for breeding, critical for the success of heterosis. This work has mapped a QTL to a gene named NLD, which encodes a phospholipase. Maize inducer lines carry a mutation in the last exon of this gene, and can be complemented by a wild type construct. GUS and fluorescent protein fusions are used to show that NLD is expressed only in mature pollen and in the pollen tube, where is localizes to the sperm cell plasma membrane.

This is a particularly exciting discovery, because it could be used to induce haploids, and consequently inbred formation in other crops, as well as to improve the process in maize. Although the mechanism is completely obscure, the significance and historical value warrants publication.

We thank the reviewer for the assessment of our manuscript and the positive feedback on our research.

Overall the paper is very well written, with just a few typos.

It is somewhat unfortunate that the QTL already has two different names, but now with the cloning is gaining a third name-and yet a fourth and fifth name with the recently published Nature and Molecular Plant papers.

The situation is indeed complex due to the history of this QTL. It should be distinguished between QTL names (*gim1* and *qhir1*) that make reference to a genetic locus and gene names that refer to a single gene and take into account gene function. By the way, our name *gim1* is anterior to the alternative appellation *qhir1*, and yet we chose the formula *gim1/qhir1* to give credit to the work of our competitors.

To clarify the situation and inform readers of the different names that have been attributed to the same gene, we now mention at the first occurrence of *NLD* in the text the other two names of GRMZM2G471240 (*MATRILINEAL* and *ZmPLA1*) and cite the corresponding recent papers.

In figure 1, the schematic representations of the mutant protein is misleading, the domain before and after the patatin domain is mostly identical, but shown as a different color, also the gray and dark green boxes are not so obvious to me.

We agree with the reviewer and changed the figure accordingly. The mostly identical domains have now the same grey shade. In addition the contrast between the three boxes within the patatin domain has been accentuated to avoid any confusion.

Figure 3, the plasma membrane localization is not clear, and could be verified by plasmolysis. Also it should be backed up by an analysis of the protein sequence, does it have a signal peptide? Is there a C-terminal membrane anchoring motif?

We agree with the reviewer that co-localization studies are indeed needed to firmly conclude on this point. Consequently the conclusions regarding the plasma membrane localization in sperm cells were tuned down (see page 9). NLD has no predicted signal peptide, but three post-translational modifications adding lipid anchors are predicted (see page 9-10). Appendix Figure S6 was added to illustrate these predictions. Please note that the last lipid anchor is lost in the truncated protein of inducing line PK6, which could explain the loss of plasma membrane localization.

Page 4, high homology does not make sense, homology is absolute, not a relative term

"Homology" was replaced by "similarity".

Page 5, the word densification should be translated into standard English

"Marker densification" was replaced by "the use of additional markers".

Referee #2:

This manuscript describes the cloning of the maize NLD gene, together with very preliminary functional characterization of the gene and protein. The identification by fine mapping and transgenic complementation is convincing, although I would have liked to see complementation with the genomic construct rather than overexpression. I am not a cell biology expert, but the protein localization data in Arabidopsis and maize look reasonable as well, with the caveat that functionality of the construct through complementation of the mutant was not confirmed.

We thank the reviewer for the assessment of our manuscript and the positive feedback on our research. We agree with his remark on overexpression, but at the time the complementation constructs were made, we had not yet validated the *NLD* promoter fragment. Complementation with a full genomic construct is presently in progress but the results will not be available for another year. Similarly, complementation with NLD::citrine fusion constructs is in progress.

The data in this manuscript are at a similar level as the recent Nature paper by Kelliher et al. While that paper included addition RNA-seq data, they were not very informative and did not add much to that paper.

Overall, both papers are of a level that one would expect in Genetics, G3, Heredity, New Phytologist, given that they do not do much more than report cloning of a mutant gene. Compared to the standard in the field, especially the raft of rice QTL cloning papers over the past few years, many of them from Chinese colleagues, papers that typically had very extensive and impressive functional data, neither the Nature paper nor the EMBO manuscript can compete. I assume Nature published the other work, despite its preliminary nature, because of potential biotechnological application. While these papers are not the first to claim that they provide a new inroad to haploid induction, the simple genetic architecture of the system makes it not unlikely that this will indeed be more widely applicable than previous technologies, with the caveat that genetic background might

influence the efficacy of haploid induction.

The current manuscript is obviously very rushed, but I suspect that the authors, to their credit, were holding out for more functional data. It seems that the authors felt merely compelled by the weak Nature paper to push this out.

Minor comments: The details on fine mapping etc. do not belong in the main text.

We feel that the positional cloning and in particular the unprecedented precision of the fine mapping (left breakpoint within the causal gene upstream of the mutation) are an integral part of the results identifying NLD as the causal gene for haploid induction in maize. The 20 lines of the already very condensed chapter seem necessary to allow the reader to appreciate the quality of these results and to guide him through the appendix data.

However, if the editor wishes to move the entire chapter or part thereof to the supplementary data, we will follow her decision.

2nd Editorial Decision 09 February 2017

Thank you for submitting the revised version of your manuscript to us. I appreciate the introduced changes, and I am happy to accept your manuscript for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: WIDIEZ Thomas
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2017-96603

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	For haploid induction rate evaluation, a minimum of 100 kernels was scored per plant p12-13
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	NA
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	NA
Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

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

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	GenBank sequence KX852318 (annotated DNA sequence of BAC clones 98E03 and 100C03 containing NLD from genotype PK6) will be released on Feb 8th
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	The manuscript contains 6 supplementary figures and 4 supplementary tables
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	Appendix Figure 1 clearly cites the sources used: Data are from Sekhon et al., 2011, The Plant Journal 66: 553-563, and Downs et al., 2013, Plant Physiology 161: 1830-1843. Tools used for sequence alignment and phylogenetic tree (Appendix Figure 2) are cited in Materials and Methods, p15. Tools used for Prediction of lipid modification sites (Appendix figure S6) are cited in Materials and Methods, P15,
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA