nature portfolio

Peer Review File

Multi-interface licensing of protein import into a phage nucleus

Corresponding Author: Dr Joseph Bondy-Denomy

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
In this paper, Kokontis et al offer the first glimpse into the phage genes and proteins involved in protein import into the nucleus-like structure of many jumbophages. This is important work in a truly untapped frontier - import/export mechanisms for an entirely new type of barrier. Here, not only do they identify, through a series of very clever selective screens, 6 "imp" genes involved in import - quite a finding in and of itself, they also show a remarkable array of specificities for Imp1 that seems unlike any protein I know of. The paper makes the following key claims; (1) that Imp1 plays an important role in protein import into the phage nucleus, (2) That Nlp1 uniquely requires Imp2 for its import, (3) That Imp1 is equipped with distinct interfaces to achieve import specificity and licensing, (4) That that 4 other loci (Imp3-6) are involved in import, the latter by direct interaction with Imp1.

There is no doubt in my mind that this is groundbreaking work, well suited to the journal. While import into membrane-bound cells/organelles is a good analogy - and finding phage proteins that do this in entirely new ways would be exciting on its own - the protein nature of the phage nucleus-like structure makes this a whole different story with potentially radically different implications (in fact I keep finding myself wondering if this system will also shed light on how some proteins are packaged into phage capsids, which is... a somewhat similar 'problem' in the field). In part out of a hopes the authors and I can disccuss this in the future (but mostly because I've been doing it for all reviews recently), I have signed the review.

The mechanism for any of this import was not revealed by the work performed. I do not feel this poses any barrier to publication - we're very much in new territory here. There is good evidence of the specificity, at the level of substrate and even at the level of residues, which I think provides a suitable scaffold for downstream inquiries outside the scope of this paper.

The writing is very strong, at some points masterful (the introduction in particular was two phenomenal paragraphs, concise and beautiful in guiding the reader to the meat of the issue).

Critical Concerns:
I have no critical concerns with the manuscript that would require a rethink or challenging of its premises

Major concerns (new experiments or analyses):
- I'm a little hesitant labelling this a major concern, because if the language (and a few diagrams) are softened a little, the manuscript is not really much weaker for it. However, it is a claim made by the manuscript that I think is not sufficiently supported. "Nlp1 uniquely requires Imp2" highlights a set of experiments that are alluded to, but missing. It's clear that Nlp1-4 require Imp1, and that Nlp1 also requires Imp2, but it is not clear that Imp2 is *not* required for Nlp2-4. Simply because the Imp2 mutations didn't show up under the Nlp2-4 selection doesn't mean they aren't yet another lower-frequency mutation (especially since not all Nlps are selecting for the same Imp1 mutations, so presumably the frequency of Imp1 mutations are not all equally likely to obscure rarer Imp2 mutations). Similarly, there's an implication in Figure 1H that Imp2 is not involved in Nlp2, 3, 4 - or in FIgure 3e that Imp 3 is not involved in Nlp2/3/4, or that IMp4,5 are not involved in Nlp1-4. But htese are never explicitly tested - and while the test using those mutant allels and trying to import an Nl1-4/Top A set of proteins wouldn't be conclusive (different residues might be involved in each), it could at least strengthen this argument considerably.

Moderate:
- Several findings make reference to the rate of escape, but this is inconsistently reported throughout (at some points not at all, e.g. "Most" L165), and given it is relied upon to justify a few experimental findings, I would prefer to see this reported more consistently, and perhaps with replicates.

Minor:
- I think I have to conceded that this ship has sailed, and that this structure will be known as a 'phage nucleus' (title, throughout) rather than (e.g. in the abstract) a "nucleus-like structure". I understand why.
- On two occasions, the authors refer to a fusion being 'inactive' (L67, LXX). It's not clear from reading whether this refers to lack of catalytic activity (EcoRI non-functional) or inability to reduce titres. The instance on L67 is clarified from a readthrough of Extended Data Fig 1A, but the second instance is not (L188) - while I'm OK with the 'data not shown" in L188, please specify exactly what is meant by "inactive".
- I wonder about the sequential naming convention for these "imp1-6" that is based on... well, the order they are presented (at least imp1 and 2 were likely discovered first) - will this come back to bite the field? Is there a cleaner way to number these?
- I think a clarification that Imp protein deletions are likely non-viable (at least, the kinds of mutations selected for suggests as much) is worth an explicit statement.
- The many interfaces of imp1 (see L135) seems to me ... remarkable. Is there an analogous 'swiss-army-knife' of a protein anywhere with that many potential substrate specificities conferred by different residues?
- I wonder if there's value in controlling the experiments about Eco localization with a methylated PhiKZ genome. The more I think about it, though, the more I think the existing experiments already suitably address the points this would alleviated.
- I did have questions about synteny and genome position very early on, that weren't touched on narratively until near the end.
- Extended Data Fig 2d-e Why is the DAPI so diffuse in the second and third rows?

Line by Line:
L13-14: I was caught off-guard by this sentence, because I missed the word "machinery" in the previous sentence. It felt like there was a imssing link between the two sentences that would better connect the two. Perhaps "while internalizing hte machinery required for DNA replication and transcription" / " required for the import of the protein machinery required for these processes, as well as the mechanism of selectivity for them remain unknown". I'm just nit-picking based on a bad reading though, feel free to ignore.
L18: likely should be preceeded by a comma
L22: the lack of specificitty as to the two queried nuclear cargos (or how many were queried) feels needlessly obfuscating.
L29: I believe Gram should be lowercase (its refering to the phylogeny, and not the Gram reaction, which would be capitalized).
L67 (or maybe L72) could start a new paragraph.
L72: "eight", L82 "one to two", L92: "eight", L216: "six" - at least by convention. I don't feel strongly about this.
L159: "not seen in" ?
L216: One reading of this is that Imp2 is required for its own import, which seems trivial :P. I'm not sure how to rephrase though, and it doesn't detract from the manuscrupt.
L224: "nearby" could specify.
References: Ref4 incorrectly capitilized journal name. Consistent failure to italicize species names in ref titles. Fair bit of alternation between title and sentence case.

Signed: Alexander Hynes

Referee #2

(Remarks to the Author)
In this manuscript, Kokontis et al. used genetic screenings to demonstrate an intriguing hierarchy of viral proteins involved during protein import into the phage nucleus. The genetic screenings are very nicely done. Overall, this is a very interesting paper with solid data to support the conclusions. Below I have a few points that need to be addressed.

Major points:

1. Fig. 1, extended data Fig. 3 and line 82.
The authors mentioned the Imp1-mNG fusion showed as puncta at the periphery of the phage nucleus. The Imp1 fusion used here was expressed from the host chromosome.
- I wonder if the expression level from the chromosomal copy is relatively similar to regular phage infections. Would it be possible that the number of puncta formed changes with the expression level of the Imp1-mNG fusion proteins?
- The authors mentioned live-cell imaging. Have the authors ever taken time-lapse movies? What time post-infection were these representative images taken? How early would the imp1 showed as localized signals?
- What's the percentage of cells with puncta? Fig. 1c only shows examples.
- Is it possible that this Imp1-mNG localization involves the interaction with the phage tubulin structure (like PhuZ spindle)? What would be special for this spot on the phage nucleus?
- The authors say Imp1-mNG is inactive, then it appears weak to claim its function.
- Is mNG-Imp2 functional?

2. Line 183 related to Imp3
The authors mentioned fluorescently tagged Imp3 did not express well, and EcoRI-Imp3 fusions were inactive, and thus the localization cannot be determined. Since Imp3 is believed to help Imp1 function properly, it will be beneficial to show at least if Imp3 is binding to or forming a complex with Imp1 or not.

3. Extended data Fig. 2
It will be helpful to show time-lapse images for transporting. What does “excluded” mean? Are the DAPI spots outside the cell phiKZ particles?

4. Fig. 2b and extended data Fig. 3
Additional information would be beneficial to clarify the result of the proposed Nlp-Imp1 interfaces:
- The authors mentioned that the PhiKZ Imp1 protein showed no significant homologs with known molecular function. How did the authors do structure-based searches? By looking at the AlphaFold predicted model, it seems that multiple domains are present in this protein. Would it be possible that slightly different AlphaFold predictions lead to an insignificant hit?
- It would be great if the authors can elaborate a little bit more about the interfaces. Does that mean these interfaces are involved in direct protein-protein bindings? By looking at the surface charges, the Imp1 mutants that escape from Imp2 fusion seem to lie on a ’tunnel’ or ‘patch’ -like positively charged surface on the predicted structure. Is it possible that this indicates an electrostatic interaction between Imp1 and Imp2? It would be beneficial to show this feature (something like a supplementary figure showing electrostatic).
- Extended data Fig. 3: why some DAPI signals are much weaker than others? Some DAPI signal only shows as small foci.

5. Extended data Fig. 4a
In 2nd and 3rd rows, are these without plaques or faint plaques?

6. Fig. 4
The authors proposed that gp67 forms a phage nuclear import complex with the protein importer Imp1. While the immunoprecipitation seems solid, it would still be preliminary without addressing the following questions:
- Is gp67 essential for ΦKZ-like phages? This could possibly explain why the genetic screening (when expressing Imp1 in trans) failed show any mutants that map to gp67.
- The stoichiometry of gp67:Imp1 seems to be 2:1. However, based on the Imp1-mNG fusion results, the actual copy number of Imp1 presented on a phage nucleus might be at a moderate level (e.g., >30 copies) while still forming only one or two puncta. Do Imp1 or Imp6 proteins by itself interact with each other as oligomers?

Minor points:
1. Line 201: Imp6-Flag should be gp67-Flag (Imp6 is only mentioned later in the text)

2. Extended data Fig. 4b was not quoted in the main text.

3. Overall, figure captions are insufficient which makes it hard to get the main conclusion from the figures.

4. Line 188: data not shown. I don’t know whether the journal allows “data not shown” in the manuscript.

Referee #3

(Remarks to the Author)
It was a great pleasure to read this new article from Dr. Joseph Bondy-Denomy’s group, which adds a new piece to the life cycle of the "phage nucleus"-forming bacteriophage. The study was conducted using classical genetic assays, during which dozens of escape mutants of the bacteriophage phiKZ were obtained and analyzed. This research identified five components of the specific protein transport system within the phage nucleus during the infection of a cell by bacteriophage phiKZ. These proteins were named Imp, with Gp69 as Imp1, Gp47 as Imp2, Gp59 as Imp3, Gp48 as Imp4, Gp287 as Imp5, and Gp67 as Imp6. Technically, the work is superbly executed. I think that Dr. Joseph Bondy-Denomy’s research significantly contributes to the understanding of phage nucleus functioning and is of substantial interest. However, previously published work from the Dr. Joe Pogliano laboratory (https://doi.org/10.1073/pnas.2321190121) and a preprint of this article (https://doi.org/10.1101/2024.03.21.585822 , posted 21 of March 2024) also revealed the function of the protein Gp69 (named PicA, Protein importer of chimalliviruses A) of bacteriophage phiKZ as a main part of the transport system. Thus, the statement from the Discussion (lines 217-219), 'A key finding from this work is the functional identification of an import factor Imp1, a previously uncharacterized protein that is broadly conserved in nucleus-forming jumbo phages,' does not seem as groundbreaking. To strengthen the paper, the authors should further explore the mechanisms of the phage nucleus transport system, both at the cellular and structural levels. On the other hand, considering the high quality of the article overall, I would recommend that the authors publish the article as it is in another journal.

Some minor questions and remarks:

1. I am not sure that it is necessary to introduce a new protein type for genes with nuclear localization, such as Nlp-proteins, instead of using their functions, like RecA-like protein Gp152 or Gp155, RNase H-like, or just retaining their annotation name for genes with unknown function. Currently, there are too many variants of important protein names in the studies of Chimaviridae phages. New ones could be confusing.

2. A fluorescent image shows Imp1 (gp69) 's position on the surface of the mature phage nucleus. However, gp69 is an early protein (as are almost all transport system components except for Imp6 (gp67)). Have you observed its colocalization with DNA during the formation of the phage nucleus?
3. Do you think one or several proteins from the Imp family could participate in the DNA transfer from the EPI vesicle into the phage nucleus?
4. In my opinion, the presented phylogenetic tree of Gp69 homologs is redundant and unrepresentative. For some genomes, coloring indicates the type of bacteria to which they belong, but for others, this coloring is not provided. This is likely because metagenomic data were extensively used in the analysis. For example, proteins QHJ76750.1 and DAO50731.1, which are marked as lacking ChmA homologs, belong to insect and human metagenomes, respectively. I recommend using complete or partial phage genomes to assess the connections between transport system components and other systems in giant bacteriophages. This analysis should be expanded and compared with the phylogenetic trees of other discovered components of the transport system

Referee #4

(Remarks to the Author)
Kokontis et al. report the identification of various phage proteins that establish a protein import network into the recently discovered “nuclei”, which are established in the interior of infected bacteria. These nuclei are formed by a network of the phage protein Chimallin or PhuN. Kokontis et al. exploited the knowledge of several phage proteins localizing into the nuclear interior (designated as Nlp1 to Nlp4) and – using an elegant genetic screen – identified several protein factors that were responsible for Nlp import. In this manner they identified a total of six import factors that enabled the import of their cargo proteins and also demonstrated a hierarchy between these. This work represents certainly a substantial and original step forward in understanding the fascinating system of phage nuclei as their defense system and is certainly interesting for a wide range of readers.

While the authors provide ample experimental evidence for the function of the detected proteins Imp1 to Imp6, a direct proof of the interaction of these “importin gene factors” with components of the nuclear shell is completely missing (see also remark below). It has been shown that Chimallin alone builds a tight protein network, which has ultrastructurally been well characterized. This network contains pores, which are too small to allow the passage of proteins. In addition, however, several proteins have been identified, which localize to the nuclear shell (e.g., reference 9). It can be assumed that one or more of these form distinct pores that allow the passage of cargo. Thus, Kokontis et al. must show the interaction of their “import genes” with these putative pore proteins, because this is the prerequisite of the import gene function. This could be achieved by a biochemical demonstration of the interaction. Especially strong, however, would be a demonstration as Frey and Görlich provided it for the interaction of import factors of eukaryotic cells and nucleoporins of the nuclear pore complex (see, e.g. Frey & Görlich, Cell 2007).

Further comments:
The first sentence on page 4 „These proteins lack…” makes a logical connection that is not justified.

Page 5, line 83: “… a localization consistent with a role in protein import.” Why actually? If Imp1 is active during protein import, I would rather expect a ring staining of the nucleus or an accumulation in the nucleus. Furthermore, the fact that the Imp1-mNeonGreen construct was inactive indicates that the results with this protein are questionable anyway.

Abbreviation mNG is not introduced.

Information on the numerical aperture defining the optical resolution of the used microscope objective lens is missing.

Ulrich Kubitscheck

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
The authors have admirably addressed my concerns. The softened language and the new cross-plating experiment both constrain the claim to what the data actually show, and further support that claim. I also appreciated the additional discussion that explained why certain aspects were kept out of the manuscript.

I have one minor remaining point about the rates of escape, see below, and a few additional comments about clarifying the magnitude of some effects in text. The rest are simply curiosity and copy-edit issues, and the authors (and editors) can disregard as their leisure – just proving I read it!

Minor Comments:
L90 “substantially decreased by this mutation” – quantify/statistics?
L107 “very low titre, however” this isn’t a very strong statement. Do you mean that plaques had low pfu counts? Or that amplifications consistently yielded low titres? Not much depends on this statement, but not entirely clear what is being quantified here.

L124: “robust phage targeting” magnitude?
L131: “lower frequency of ~10^-7.” Given the frequency stated earlier was “10^6 to 10^7” it really is hard to know how much “lower” this frequency is.

Curiosity:
L72: Import and Impart. Clever. I like it.
L155: Double mutants no apparent fitness cost?
L165: Do relevant residues co-evolve with e.g. some residues of the NIps?
L184: The genetically dominant mutations are a neat phenotype. This isn’t a question, just a statement.

Copy Editing:
L56: This sentence is awkward, as it implies the Imp2/4/5 do not have a role in protein import. Consider rewording entirely.
L76: Tense change (provides) a little awkward.
L95: mNG already introued in line 91.
L96: “and likely first initiates protein import” not clear where this is coming from (I don’t disagree that import probably starts there, but there’s neither a citation, nor experiment, nor reasoning provided”)
L112: imp2 is underlined
L115: “Imp2 only appears to be conserved” reword “imp2 appears to only be conserved”
L127: I think use of non-redundant didn’t help here, consider rewording e.g. “none of which overlapped with those previously isolated”?
L134: line break before section not present (was in prior section). Also missing before next section. L192: Why not mention the ratio of imp1 mutations to imp3 mutations, as for the earlier screen?
References still incompletely edited, e.g. 6 lacks ital, 7 lacks ital and is in title case, 7 lacks ital and is title case, 11 lacks ital and is in title case, 12 in title case, 14 lacks capitalization of pseudomonas and ital, 17 lacks ital, 18 lacks ital, 19 lacks ital,

Referee #2

(Remarks to the Author)
The authors have nicely addressed my concerns. One additional suggestion is to add time stamps onto the live-cell movie.

Referee #3

(Remarks to the Author)
I appreciate your thorough responses to my questions. The clear emphasis on the specificity of the transport system into the phage nucleus as the main finding of the study highlights the significance of the meticulous work you have conducted. I am pleased to recommend the manuscript for publication. Please address the following two minor points before finalizing:

1.The paragraph lines 68-78 could be clearer. The current wording may be confusing as it first mentions four fused nuclear proteins and only later clarifies that these were selected from a larger pool. It would be beneficial to rephrase this section to explicitly state that a broader set of NIps was initially tested, from which only four showed a significant reduction in phage titer. This adjustment would help clarify the sequence of events for the reader.

2. A minor point regarding Figure 2: Please add information about which phages were cultured in the figure description. I assume they were labeled by the colors corresponding to the bacterial species, but it would be helpful to make this explicit in the description.

Referee #4

(Remarks to the Author)
Here, I will only respond to the rebuttal letter and the additional experiments, which have been done in response to my original comments.

The direct evidence of binding between Imp1 and its cargo Nlp2 certainly validates the overall approach used in this work, which is further supported by the results with the various mutants mentioned.

However, in my first review I was not referring to the binding between Imp1 and its cargo, but to the binding between Imp1 and the components of the shell. Enustun et al. (2023) (ref 9) identified six proteins that associate with the nuclear shell, including gp2, gp61, gp63 and gp64. Demonstrating the direct interaction of imp1 with one (or more) of these proteins would greatly strengthen the manuscript. More importantly, the identification of an interaction between a shell protein and imp1 would give a first indication of what kind of selection mechanism is present and how directionality is achieved in this intriguing system.

Ulrich Kubitscheck

Version 2:

Reviewer comments:

Referee #4

(Remarks to the Author)
I thank the authors for their detailed and clarifying answers to my questions. I particularly recognize that they have made considerable efforts to demonstrate a direct interaction between ChmA and Imp1, and it is certainly unsurprising that ChmA has solubility problems. Unfortunately, as the authors state the experiments performed have limited validity and I fully agree with the authors that they should not be included in the publication at this time. I am very happy to recommend this carefully conducted work for publication.

We thank the reviewers and editors for their constructive, and largely positive, comments about our manuscript. We have amended the text where suggested and added the following major experiments to bolster our claims and respond to reviewer comments:

1. Co-immunoprecipitation reveals a direct binding event between Imp1 and Nlp2, in a manner that is disrupted by Imp1 mutations that abolish Nlp2 import, but not by Imp1 mutations that abolish import of other Nlp proteins.
2. Time lapse microscopy to show Imp1 localizing early to the phage shortly after DNA injection, suggesting that it is assembled into the nascent phage nucleus early in the infection. We also have added (for reviewers) a time lapse of a pre-expressed protein being imported into the phage nucleus to address a couple of questions.
3. Updated text and table to reflect that phage escape mutants isolated under one particular pressure do not stop importing other cargo. This was shown previously but we have expanded these observations with more ‘cross-plating’ of mutant phages against other EcoRI-fusion proteins. The take home message remains that there is quite a lot of specificity, where one mutation abolishes the import of one cargo but not others.

Referees' comments:

Referee #1 (Remarks to the Author):

In this paper, Kokontis et al offer the first glimpse into the phage genes and proteins involved in protein import into the nucleus-like structure of many jumbophages. This is important work in a truly untapped frontier - import/export mechanisms for an entirely new type of barrier. Here, not only do they identify, through a series of very clever selective screens, 6 "imp" genes involved in import - quite a finding in and of itself, they also show a remarkable array of specificities for Imp1 that seems unlike any protein I know of. The paper makes the following key claims; (1) that Imp1 plays an important role in protein import into the phage nucleus, (2) That Nlp1 uniquely requires Imp2 for its import, (3) That Imp1 is equipped with distinct interfaces to achieve import specificity and licensing, (4) That that 4 other loci (Imp3-6) are involved in import, the latter by direct interaction with Imp1.

There is no doubt in my mind that this is groundbreaking work, well suited to the journal. While import into membrane-bound cells/organelles is a good analogy - and finding phage proteins that do this in entirely new ways would be exciting on its own - the protein nature of the phage nucleus-like structure makes this a whole different story with potentially radically different implications (in fact I keep finding myself wondering if this system will also shed light on how some proteins are packaged into phage capsids, which is... a somewhat similar 'problem' in the field). In part out of a hopes the authors and I can disccuss this in the future (but mostly because I've been doing it for all reviews recently), I have signed the review.

The mechanism for any of this import was not revealed by the work performed. I do not feel this poses any barrier to publication - we're very much in new territory here. There is good evidence of the specificity, at the level of substrate and even at the level of residues, which I think provides a suitable scaffold for downstream inquiries outside the scope of this paper.

The writing is very strong, at some points masterful (the introduction in particular was two phenomenal paragraphs, concise and beautiful in guiding the reader to the meat of the issue).

Critical Concerns:
I have no critical concerns with the manuscript that would require a rethink or challenging of its premises

Thank you for your glowingly positive comments about our work.

Major concerns (new experiments or analyses):
- I'm a little hesitant labelling this a major concern, because if the language (and a few diagrams) are softened a little, the manuscript is not really much weaker for it. However, it is a

claim made by the manuscript that I think is not sufficiently supported. "Nlp1 uniquely requires Imp2" highlights a set of experiments that are alluded to, but missing. It's clear that Nlp1-4 require Imp1, and that Nlp1 also requires Imp2, but it is not clear that Imp2 is *not* required for Nlp2-4. Simply because the Imp2 mutations didn't show up under the Nlp2-4 selection doesn't mean they aren't yet another lower-frequency mutation (especially since not all Nlps are selecting for the same Imp1 mutations, so presumably the frequency of Imp1 mutations are not all equally likely to obscure rarer Imp2 mutations).

We thank the reviewer for this point. Indeed, we agree that we can not be sure that Imp2 is not required for import of other factors, but simply Imp2 mutations did not arise with any other selection. We have removed the word 'uniquely' from the offending sentence (Line 183), now reading "...these data demonstrate that Nlp1 requires Imp2 for its import, …". Earlier in the same paragraph, we have maintained this word as follows "Imp2 mutations were uniquely selected for by EcoRI-Nlp1".

Similarly, there's an implication in Figure 1H that Imp2 is not involved in Nlp2, 3, 4 - or in FIgure 3e that Imp 3 is not involved in Nlp2/3/4, or that IMp4,5 are not involved in Nlp1-4. But these are never explicitly tested - and while the test using those mutant alleles and trying to import an Nl1-4/Top A set of proteins wouldn't be conclusive (different residues might be involved in each), it could at least strengthen this argument considerably.

We thank the reviewer for this comment and agree with the reviewer's conclusions. As this reviewer suggested, we have performed further "cross-plating" experiments to address whether *imp2* mutant phages import and are restricted by EcoRI fusions to Nlp2-4. Results show that all *imp2* mutant phages maintain import of all other constructs. This has been added to the manuscript as Extended data figure 1d, to go with this statement (L146):
"*imp2* mutations were uniquely selected for by EcoRI-Nlp1 (but not by EcoRI-Nlp2/3/4) and *imp2* mutant phages still imported and succumbed to EcoRI fusions to Nlp2, Nlp3, or Nlp4, demonstrating that these mutations specifically perturb Nlp1 import"

Moderate:
- Several findings make reference to the rate of escape, but this is inconsistently reported throughout (at some points not at all, e.g. "Most" L165), and given it is relied upon to justify a few experimental findings, I would prefer to see this reported more consistently, and perhaps with replicates.
This has been clarified to a specific number in each case that we could find and the word "most" no longer appears in the manuscript.

Minor:
- I think I have to conceded that this ship has sailed, and that this structure will be known as a 'phage nucleus' (title, throughout) rather than (e.g. in the abstract) a "nucleus-like structure". I understand why.
Indeed this has been somewhat inconsistent in the field but we have settled on phage nucleus. We thank the reviewer for understanding.

- On two occasions, the authors refer to a fusion being 'inactive' (L67, LXX). It's not clear from reading whether this refers to lack of catalytic activity (EcoRI non-functional) or inability to reduce titres. The instance on L67 is clarified from a readthrough of Extended Data Fig 1A, but the second instance is not (L188) - while I'm OK with the 'data not shown" in L188, please specify exactly what is meant by "inactive".
We have clarified instances of this phrase in the text at each line number mentioned by the reviewer.

- I wonder about the sequential naming convention for these "imp1-6" that is based on... well, the order they are presented (at least imp1 and 2 were likely discovered first) - will this come back to bite the field? Is there a cleaner way to number these?
Thank you for this comment. We have indeed gone through many iterations of naming these proteins, including simply referring to them as 'gp' names or where their function is known (as suggested by reviewer 3). After trying different options and sharing our manuscript with colleagues we settled on the somewhat old-fashioned methods from genetic screens that identify mutants based on a phenotype that was selected for (i.e. *sec* mutants from Schekman, etc), generally in order of presentation/discovery, as opposed to modern operon-based naming of orf1-orf2-orf3. Although we conceded that alternatives may also work, we have opted to keep with this nomenclature, if agreeable.

- I think a clarification that Imp protein deletions are likely non-viable (at least, the kinds of mutations selected for suggests as much) is worth an explicit statement.
We appreciate this reviewer's comment where they are essentially 'reading between the lines' and agree with this sentiment. We initially opted to not get into this but we have now added text in results (L188):
"This mutational spectrum, with no nonsense mutations or frameshifts identified, also suggests that *imp1* is essential."

We were also unable to knockout *imp1* or *imp6* with our knockout method (which works well for non-essential genes), but opted to not add this as simply failing to make a knockout could be argued to be inconclusive. Therefore, we will lean on the fact that we have >30 mutant alleles of *imp1,* with none appearing to be nulls.

- The many interfaces of imp1 (see L135) seems to me ... remarkable. Is there an analogous 'swiss-army-knife' of a protein anywhere with that many potential substrate specificities conferred by different residues?
We also think this is a really important and exciting aspect of our paper.
After consulting with colleagues and literature, it seems that there are many examples of this in the literature, including Ras with binding sites for different proteins. WD-40 domains and chaperones like Hsp90 have many binding sites for different clients, as do co-chaperones. We have opted not to elaborate on this in the discussion in the interest of staying focused and not attempting to promote other aspects of novelty.

- I wonder if there's value in controlling the experiments about Eco localization with a methylated PhiKZ genome. The more I think about it, though, the more I think the existing experiments already suitably address the points this would alleviated.
We agree. Dead EcoRI constructs have been our preferred control throughout, as they can be used with a single phage stock and is arguably more subtle perturbation to the experimental system than 92 methylation events which could have unknown consequences on phage transcription, DNA replication, etc.

- I did have questions about synteny and genome position very early on, that weren't touched on narratively until near the end.
This was a topic that was hard to address earlier in the paper as the topic builds, but is one reason we present the orf numbers and orf maps before renaming them, so the reader can track along the genomic localization (or lack of neighborhood). We also draw the reviewer's attention to the neighborhood orf maps that are included in every figure as new genes are introduced.

- Extended Data Fig 2d-e Why is the DAPI so diffuse in the second and third rows?
The import of an EcoRI construct during a WT phage infection leads to faint phage nucleus DAPI staining, due to phage DNA being recognized/cut by EcoRI import and replication is halted. We also occassionally see 'health' phage nuclei that do not stain well with DAPI for reasons unknown, but in each case, we have a red imported/excluded protein that can also be used to see where the phage nucleus is and know that it is formed in that cell.

Line by Line:
L13-14: I was caught off-guard by this sentence, because I missed the word "machinery" in the previous sentence. It felt like there was a missing link between the two sentences that would better connect the two. Perhaps "while internalizing the machinery required for DNA replication and transcription" / " required for the import of the protein machinery required for these processes, as well as the mechanism of selectivity for them remain unknown". I'm just nit-picking based on a bad reading though, feel free to ignore.
OK

L18: likely should be preceded by a comma
Fixed. L22 now.

L22: the lack of specificity as to the two queried nuclear cargos (or how many were queried) feels needlessly obfuscating.
The two cargos are now specified

L29: I believe Gram should be lowercase (its referring to the phylogeny, and not the Gram reaction, which would be capitalized).
Fixed

L67 (or maybe L72) could start a new paragraph.
Fixed. New paragraph at L67, thank you

L72: "eight", L82 "one to two", L92: "eight", L216: "six" - at least by convention. I don't feel strongly about this.
Corrected

L159: "not seen in" ?
Fixed

L216: One reading of this is that Imp2 is required for its own import, which seems trivial :P. I'm not sure how to rephrase though, and it doesn't detract from the manuscript.
This is a good point. The offending sentence has been edited to say:
"we have identified five factors (Imp1-5) required, in different combinations, for the import of six6 proteins (Nlp1-4, Imp2, TopA) (Table1)." We hope that referring to the Table will clarify which mutants come up under which selections.

L224: "nearby" could specify.
Changed to "in the same locus"

References: Ref4 incorrectly capitalized journal name. Consistent failure to italicize species names in ref titles. Fair bit of alternation between title and sentence case.
Fixed

Signed: Alexander Hynes

Referee #2 (Remarks to the Author):

In this manuscript, Kokontis et al. used genetic screenings to demonstrate an intriguing hierarchy of viral proteins involved during protein import into the phage nucleus. The genetic screenings are very nicely done. Overall, this is a very interesting paper with solid data to support the conclusions. Below I have a few points that need to be addressed.

Major points:

1. Fig. 1, extended data Fig. 3 and line 82.
The authors mentioned the Imp1-mNG fusion showed as puncta at the periphery of the phage nucleus. The Imp1 fusion used here was expressed from the host chromosome.
- I wonder if the expression level from the chromosomal copy is relatively similar to regular phage infections. Would it be possible that the number of puncta formed changes with the expression level of the Imp1-mNG fusion proteins?
Indeed we agree that it is possible that the number of puncta can increase with expression level. When we expressed Imp-mNG from a high copy plasmid in the cell, we often see two or three puncta when we induce the plasmid. We have opted to not add this to the paper however as we typically use chromosomal insertion to keep expression levels reasonable.

- The authors mentioned live-cell imaging. Have the authors ever taken time-lapse movies?
Thank you for this suggestion. We have added a time lapse image of Imp1 during infection and think it is exciting and supports that Imp1 is localizing with the nascent phage nucleus as it is assembling, near the polar site of infection. (Supplementary Video 1 now with time lapse images in Extended Data Fig. 3b). Text on L117.

What time post-infection were these representative images taken?
Imaging at 50-60 min post infection, clarified in Figure legend, L462.

How early would the imp1 showed as localized signals?
Imp1 localizes to the nascent nucleus near the pole at ~15 minutes post infection, this is added in time lapse as mentioned above. This focus then moves with the phage nucleus.

- What's the percentage of cells with puncta? Fig. 1c only shows examples.
100% (154/154) infected cells have puncta, added to the results section and figure legend. L117 and L462.

- Is it possible that this Imp1-mNG localization involves the interaction with the phage tubulin structure (like PhuZ spindle)? What would be special for this spot on the phage nucleus?
The PhuZ spindle is a dynamic structure that builds and tears down rapidly, pushing and rotating the phage nucleus. So we do not think this is a good candidate for a phage protein that would mark a single spot on the phage nucleus.

- The authors say Imp1-mNG is inactive, then it appears weak to claim its function.

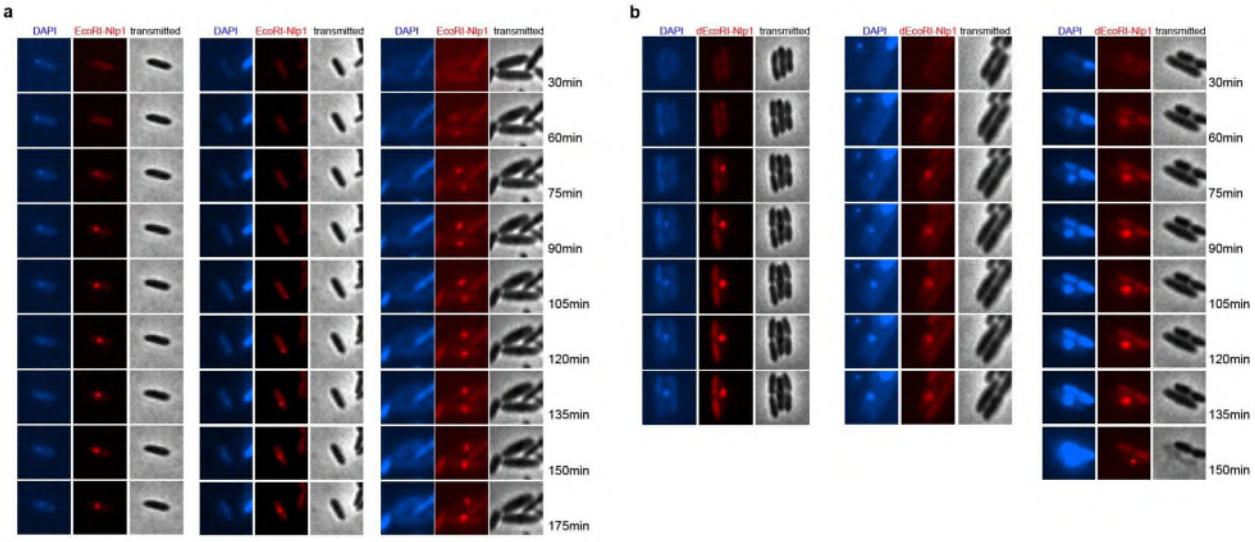
To clarify, the function is claimed based on the mutations isolated under selection with different imported cargo, including plaque assays, microscopy, and complementation studies. Those experiments rely on naturally expressed Imp1, encoded by the phage, which is the strength of our genetic selection approach.
The localization is the only item we base on this fusion protein, which is unfortunately non-functional for complementation. To address the reviewer comment, we tagged Imp1 with a small myc tag instead (which did *not* hinder function of Imp1) for use in immunofluorescence (IF). This posed its own technical challenges, however, as three different anti-myc antibodies (Thermo, Santa Cruz, Cell Signaling) attempted showed non-specific staining during phage infection. We therefore think it is best to show the results of localization (and refer to previous papers that have shown something similar for the homologous protein in other phages), but be transparent about the tag interfering with some aspect of import function. We also note that other groups in this field routinely use over-expressed fluorescent fusions of phage proteins (including a competing paper on this topic) but have never assessed whether they complement a mutant and thus are not aware.

- Is mNG-Imp2 functional?
This fusion is also non-functional for complementation of an Imp2 mutant. The functional role of Imp2 is again derived not from this fusion but from mutations in the phage genome that abolish the import of Nlp1. The localization of Imp2 inside the nucleus is orthogonally supported by its fusion to EcoRI which leads to phage DNA targeting, while EcoRI alone does not.

2. Line 183 related to Imp3
The authors mentioned fluorescently tagged Imp3 did not express well, and EcoRI-Imp3 fusions were inactive, and thus the localization cannot be determined. Since Imp3 is believed to help Imp1 function properly, it will be beneficial to show at least if Imp3 is binding to or forming a complex with Imp1 or not.
We agree in general with this suggestion but efforts to express Imp3 have remained unsuccessful. We think the value of the genetic result is still high as it implicates this locus/operon, which has a highly conserved gene (imp3), and is in an operon that has not previously been implicated in any function in the literature. We have however added new genetic data that this reviewer may find interesting (and should support future work on this gene). Please refer to Line 300 where we have strengthened the genetic data around Imp3, demonstrating that it likely works in concert with Imp1, Imp4, and Imp5 to facilitate import of host TopA. Perhaps interactions between some of these factors to be elucidated in future work will solve the current Imp3 expression challenges.

3. Extended data Fig. 2
It will be helpful to show time-lapse images for transporting. What does “excluded” mean? Are the DAPI spots outside the cell phiKZ particles?
Thank you for this suggestion. The figure legend (L757) has been amended to include our definition of “excluded,” where it now reads: “‘Excluded refers to localization of sfCherry2-fused proteins outside of the phage nucleus.” DAPI spots outside the cell are ΦKZ particles.

As requested by this reviewer, we include here examples of time lapse imaging of WT phage infections with EcoRI-sfCherry2-Nlp1 or catalytic dead dEcoRI-sfCherry2-Nlp1. Phage nuclei that can form in cells expressing active EcoRI-sfCherry2-Nlp1 often remain small and faintly DAPI stained, with Cherry labeled protein inside, while dead EcoRI construct localizes Cherry-labeled protein and grows large with strong DAPI stain. We provide those here but have not changed the primary figures as the Nlp1 time lapse did not reveal any transient "build up" at an Imp1 focus, nor did they reveal anything notable other than the accumulation of protein. We include these images here for reviewers but have not added them to the paper as they are somewhat redundant with what we previously published in Mendoza et al. 2020 *Nature (*Fig 3E in that paper).
Overall these experiments bolster our belief that the images we have shown of included/excluded protein in the paper do not leave anything out that exhaustive time lapse imaging might reveal.



4. Fig. 2b and extended data Fig. 3
Additional information would be beneficial to clarify the result of the proposed Nlp-Imp1 interfaces:
- The authors mentioned that the PhiKZ Imp1 protein showed no significant homologs with known molecular function. How did the authors do structure-based searches?
An AlphaFold2 model was generated and then queried using DALI. Methods have been updated to clarify this. L625.
By looking at the AlphaFold predicted model, it seems that multiple domains are present in this protein. Would it be possible that slightly different AlphaFold predictions lead to an insignificant hit?
We are not exactly sure what is meant by this question, but we can clarify that the AlphaFold prediction score was very strong across the whole protein, even if different domains may be present.

- It would be great if the authors can elaborate a little bit more about the interfaces. Does that

mean these interfaces are involved in direct protein-protein bindings? By looking at the surface charges, the Imp1 mutants that escape from Imp2 fusion seem to lie on a 'tunnel' or 'patch' -like positively charged surface on the predicted structure. Is it possible that this indicates an electrostatic interaction between Imp1 and Imp2? It would be beneficial to show this feature (something like a supplementary figure showing electrostatic).
Great observation. We thank the reviewer for this suggestion and have added the electrostatic map to Extended Data Fig. 4b. Not all interfaces have strong negative charge, but that one certainly does. While we have not demonstrated that Imp1 directly binds to each cargo, this certainly one model that we outline in the discussion. We have indeed added exciting new data (Fig. 4e) showing binding between Imp1 and Nlp2, which is broken by specific mutations in Imp1.

- Extended data Fig. 3: why some DAPI signals are much weaker than others? Some DAPI signal only shows as small foci.
As mentioned above in response to another question from this reviewer, we have clarified this with time lapse images shown here, to clarify that if active EcoRI is imported into the phage nucleus, we can still see the nucleus form, with EcoRI protein localized inside, but DAPI staining is often detectable but faint. We often see the appearance of a phage nucleus 'attempting' to form but never quite getting to full DNA replication due to EcoRI localization. There are also occassionally cells where DAPI staining is faint for unknown reasons, but we direct the reviewer to the red fluoresecent protein that is always shown in parallel where it is clear that the phage nucleus is formed and where it is.

5. Extended data Fig. 4a
In 2nd and 3rd rows, are these without plaques or faint plaques?
There are no plaques here in rows 2 and 3. The associated text (L233) described Imp1 mutants fused to EcoRI that "maintained full phage targeting with no escape mutants isolated"

6. Fig. 4
The authors proposed that gp67 forms a phage nuclear import complex with the protein importer Imp1. While the immunoprecipitation seems solid, it would still be preliminary without addressing the following questions:
- Is gp67 essential for ΦKZ-like phages? This could possibly explain why the genetic screening (when expressing Imp1 in trans) failed show any mutants that map to gp67.
A recent paper from the Corbett lab (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11077065/) reported a gp67 (ChmC in the paper) homolog is essential in a related phage infecting *E. coli*. Its knockdown prevented plaque formation. A very recent preprint from the Vogel lab also suggests that gp67 (Imp6) and gp69 (Imp1) are essential based on mRNA knockdown experiments. We have added text in the result (L189) that clarifies that Imp1 is likely essential (no mutants isolated were stop or frameshift mutations). We have also referenced the Corbett paper to this effect re: Imp6/gp67.

We were also unable to knockout *imp1* or *imp6* with our knockout method (which works well for non-essential genes), but opted to not add this as simply failing to make a knockout could be

argued to be inconclusive. Therefore, we will lean on the fact that we have >30 mutant alleles of *imp1,* with none appearing to be nulls. On this point, we do think it is possible for mutations to arise in essential genes (as they have for *imp1*) which is a strength of this approach. Therefore, for now, it is unclear why mutations did not map to gp67 but perhaps gp69 is more mutable due to the many cargo specific interfaces.

- The stoichiometry of gp67:Imp1 seems to be 2:1. However, based on the Imp1-mNG fusion results, the actual copy number of Imp1 presented on a phage nucleus might be at a moderate level (e.g., >30 copies) while still forming only one or two puncta. Do Imp1 or Imp6 proteins by itself interact with each other as oligomers?
We were also surprised by this result, expecting to at least see Imp1 form oligomers. No, Imp1 by itself does not oligomerize, shown in the SEC trace in Figure 4C. The absence of Imp1 oligomerization *in vitro* actually motivated our search for interactors reasoning that a larger complex could emerge with the right partners. We have not tested Imp6 oligomerization alone but the paper mentioned above on ChmC (gp67 homolog) showed that it forms a tetramer in solution. We suspect that future work with factors identified in this paper and in the future will reveal a larger complex assembly.

Minor points:
1. Line 201: Imp6-Flag should be gp67-Flag (Imp6 is only mentioned later in the text)
Corrected, thank you. L315.
2. Extended data Fig. 4b was not quoted in the main text.
Corrected, thank you. We now refer to it as new Imp1 mutations arise from various selections.
3. Overall, figure captions are insufficient which makes it hard to get the main conclusion from the figures.
The figure caption (bolded statement at start of figure) has been updated for all four main figures to hopefully enhance the main conclusion understanding.
4. Line 188: data not shown. I don’t know whether the journal allows “data not shown” in the manuscript.
Removed

Referee #3 (Remarks to the Author):

It was a great pleasure to read this new article from Dr. Joseph Bondy-Denomy’s group, which adds a new piece to the life cycle of the "phage nucleus"-forming bacteriophage. The study was conducted using classical genetic assays, during which dozens of escape mutants of the bacteriophage phiKZ were obtained and analyzed. This research identified five components of the specific protein transport system within the phage nucleus during the infection of a cell by bacteriophage phiKZ. These proteins were named Imp, with Gp69 as Imp1, Gp47 as Imp2, Gp59 as Imp3, Gp48 as Imp4, Gp287 as Imp5, and Gp67 as Imp6. Technically, the work is superbly executed. I think that Dr. Joseph Bondy-Denomy’s research significantly contributes to the understanding of phage nucleus functioning and is of substantial interest. However, previously published work from the Dr. Joe Pogliano laboratory

(https://doi.org/10.1073/pnas.2321190121) and a preprint of this article (https://doi.org/10.1101/2024.03.21.585822 , posted 21 of March 2024) also revealed the function of the protein Gp69 (named PicA, Protein importer of chimalliviruses A) of bacteriophage phiKZ as a main part of the transport system. Thus, the statement from the Discussion (lines 217-219), 'A key finding from this work is the functional identification of an import factor Imp1, a previously uncharacterized protein that is broadly conserved in nucleus-forming jumbo phages,' does not seem as groundbreaking. To strengthen the paper, the authors should further explore the mechanisms of the phage nucleus transport system, both at the cellular and structural levels. On the other hand, considering the high quality of the article overall, I would recommend that the authors publish the article as it is in another journal.

Thank you for these comments about the strength of the paper.
1. We point out that although the mentioned paper also identified gp69, that work revealed only a single interface with 5 mutations that impact only one cargo. They were limited by using a nuclease with a single cut site and thus often had uninformative mutations. Regardless, the mentioned work does not reveal the key role of gp69 as a *specificity* factor, which is where we focus, in addition to our unique discovery of *imp2, imp3, imp4, imp5* and and now two direct interactors with Imp1 (Imp6 and Nlp2). We therefore would respectfully suggest that our work (which was preprinted at the same time) provides a significant advance.

We have also amended the offending sentence to reflect our focus on specificity (L370):
"A key finding from this work is the functional identification of an import and specificity factor Imp1, a previously uncharacterized protein that is broadly conserved in nucleus-forming jumbo phages"

2. In response to the comment about how we can strengthen the paper, we have indeed added time lapse images to show the early localization of Imp1 to the infected pole and demonstrated a direct interaction between Imp1 and a cargo protein, which can be destabilized by specific mutations in Imp1. We consider that this result will provide a clear road map to future studies of cargo-import specificity.

Some minor questions and remarks:

1. I am not sure that it is necessary to introduce a new protein type for genes with nuclear localization, such as Nlp-proteins, instead of using their functions, like RecA-like protein Gp152 or Gp155, RNase H-like, or just retaining their annotation name for genes with unknown function. Currently, there are too many variants of important protein names in the studies of Chimaviridae phages. New ones could be confusing.
Thank you for this comment. We have indeed gone through many iterations of naming these proteins, including simply referring to them as 'gp' names or where their function is known. While some proteins like gp152 have a recA-like domain, other imported proteins do not have a predicted function and the newly implicated *"imp"* genes have no known domains. We were therefore concerned that having a paper discuss >10 proteins with 'gpXX' instead of standardizing the nomenclature with Nlp and Imp factors would be very confusing. Although we

conceded that alternatives may also work, we have opted to keep with this nomenclature. Hopefully this is not so offensive that it damages the paper.

2. A fluorescent image shows Imp1 (gp69) 's position on the surface of the mature phage nucleus. However, gp69 is an early protein (as are almost all transport system components except for Imp6 (gp67)). Have you observed its colocalization with DNA during the formation of the phage nucleus?
We have added time lapse images which indeed shows what the reviewer is suggesting, that gp69/Imp1 localizes early in the infection closer to the pole where the phage first injects its DNA and starts to build the protein nucleus. We do think that gp69/Imp1 is required early as the nucleus is just being built. These exciting data are now presented in Extended Data Fig. 3b and L117 and Supplementary Video 1.

3. Do you think one or several proteins from the Imp family could participate in the DNA transfer from the EPI vesicle into the phage nucleus?
We do not think that this is a *direct* role for the proteins described here. We suspect that proteins in the EPI vesicle that might play a role are injected proteins and factors in the nascent nucleus would be localized there by the Imp system, but not be the Imp proteins themselves. This is speculation though, and since we did not design a selection around this function we have no reason to suggest additional roles for the Imp proteins. As a lab we are very interested in this DNA transfer event, but have not speculated in this paper.

4. In my opinion, the presented phylogenetic tree of Gp69 homologs is redundant and unrepresentative. For some genomes, coloring indicates the type of bacteria to which they belong, but for others, this coloring is not provided. This is likely because metagenomic data were extensively used in the analysis. For example, proteins QHJ76750.1 and DAO50731.1, which are marked as lacking ChmA homologs, belong to insect and human metagenomes, respectively. I recommend using complete or partial phage genomes to assess the connections between transport system components and other systems in giant bacteriophages. This analysis should be expanded and compared with the phylogenetic trees of other discovered components of the transport system
The reviewer is correct. We labeled bacteria where phage isolates had been cultured but did not label metagenomes. If the reviewer is suggesting removing anything that derived from metagenomes, we respectfully think that including all identified homologs is an asset, not a downside. The color shading is therefore included to help the reader see which homologs are in isolated phages, but we still harness the sequence information from metagenomes as they remain “real” proteins.

Regarding it being “redundant”, we interpret this to mean that the reviewer thinks that the protein sequences are redundant (i.e. identical/near-identical sequences) from the metagenomic data, but that does not appear to be the case. Many different clades are present within the large segment of the tree that derives from non-isolated phages which therefore seems like an asset here and precisely why metagenomics is done.

To address the second point about other import factors, we have added a new figure (Extended Data Fig. 5b) that highlights the co-occurrence for Imp1-6, ChmA across representative jumbophages, showing conservation of Imp1, 3, 6 and lower conservation of others. This is largely adapted from Prichard et al. Cell Reports, a paper from the Pogliano lab that compiled a helpful list of conserved genes in this family, within diverse phages and those that are often used as experimental models.

Referee #4 (Remarks to the Author):

Kokontis et al. report the identification of various phage proteins that establish a protein import network into the recently discovered “nuclei”, which are established in the interior of infected bacteria. These nuclei are formed by a network of the phage protein Chimallin or PhuN. Kokontis et al. exploited the knowledge of several phage proteins localizing into the nuclear interior (designated as Nlp1 to Nlp4) and – using an elegant genetic screen – identified several protein factors that were responsible for Nlp import. In this manner they identified a total of six import factors that enabled the import of their cargo proteins and also demonstrated a hierarchy between these. This work represents certainly a substantial and original step forward in understanding the fascinating system of phage nuclei as their defense system and is certainly interesting for a wide range of readers.

While the authors provide ample experimental evidence for the function of the detected proteins Imp1 to Imp6, a direct proof of the interaction of these “importin gene factors” with components of the nuclear shell is completely missing (see also remark below). It has been shown that Chimallin alone builds a tight protein network, which has ultrastructurally been well characterized. This network contains pores, which are too small to allow the passage of proteins. In addition, however, several proteins have been identified, which localize to the nuclear shell (e.g., reference 9). It can be assumed that one or more of these form distinct pores that allow the passage of cargo. Thus, Kokontis et al. must show the interaction of their “import genes” with these putative pore proteins, because this is the prerequisite of the import gene function. This could be achieved by a biochemical demonstration of the interaction. Especially strong, however, would be a demonstration as Frey and Görlich provided it for the interaction of import factors of eukaryotic cells and nucleoporins of the nuclear pore complex (see, e.g. Frey & Görlich, Cell 2007).

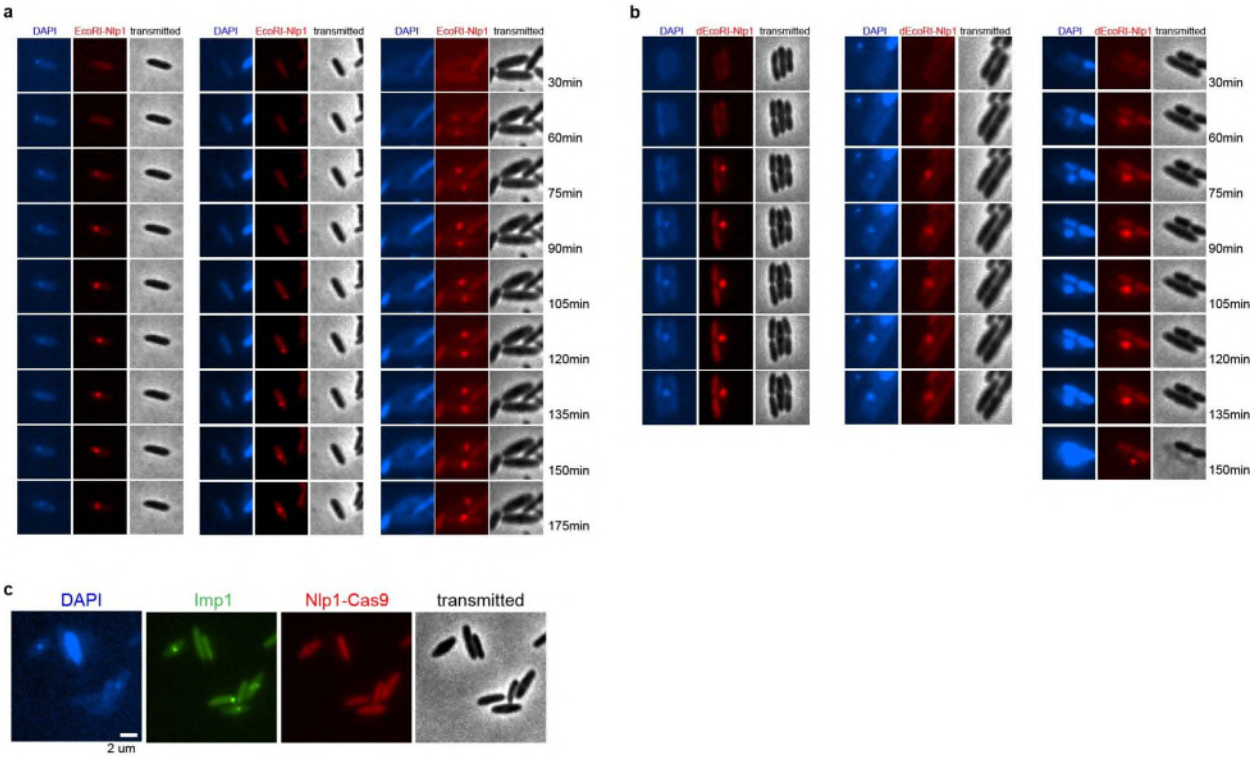
We thank this reviewer for their strong comments about the work.

To address the short-coming described, we took two approaches, both suggested by the reviewer:

- First we used *in vitro* biochemistry to test whether the Nlp cargo proteins interact directly with Imp1. Excitingly, we have now added to the paper that gp155 (Nlp2) interacts with Imp1 in a pull-down where both proteins are co-expressed in *E. coli*. Excitingly, mutations in Imp1 (H300Y and E310G) that emerged under Nlp2-EcoRI selection abolish the pull-down result while other mutations in Imp 1 (R320S and T110N) that emerged under different selection pressures do not perturb the pulldown result. As a positive control, we show that gp67 (Imp6,

a proposed complex member with gp69/Imp1) maintains binding to the H300Y and E310G mutants that lose Nlp2 binding. This positive control (gp67 binding) supports that these Imp1 mutants are still well folded, but lose Nlp2 binding due to specific mutations in the binding site that emerged from *in vivo* selection. In addition to this specific binding result, we argue that this result validates the overall approach we have used in this manuscript. These data are presented in Fig. 4e and starting at L337.

- We next turned to microscopy to address comments from this reviewer and others, imaging the following constructs: Imp1-mNg (with time lapse), EcoRI-sfCherry2-Nlp1 (imported, with time lapse), deadEcoRI-sfCherry2-Nlp1 (imported, with time lapse), and Cas9-sfCherry2-Nlp1 (non-imported).
    - We have added time lapse images which indeed shows what another reviewer suggested, that gp69/Imp1 localizes early in the infection close to the pole where the phage first injects its DNA and starts to build the protein nucleus. These exciting data are now presented around L117, Extended Data Fig 3b, Supp Video 1.
    - EcoRI-sfCherry2-Nlp1, deadEcoRI-sfCherry2-Nlp1 appear imported into the phage nucleus over time, as expected. When imaged with Imp1 labeled, we do not see any notable accumulation with this protein over time, likely because import is too rapid as this resolution. Phage nuclei that are able to form in cells expressing active EcoRI-sfCherry2-Nlp1 often remain small and faintly DAPI stained (a, n = 87/90 remain small/faintly stained, 2 replicates), with Cherry labeled protein inside, while dead EcoRI construct localizes Cherry labeled protein and grows large with strong DAPI stain (b, n = 79/91 grow large/strong stain, 2 replicates). We provide these images here but have not changed the primary figures in the paper
    - Cas9-sfCherry2-Nlp1 was used to attempt to see a build up or co-localization with Imp1-mNg as it is not imported successfully, presumably because it is too big (first shown in Mendoza *Nature* 2020). However, we simply observed exclusion, with no build up at 60 min post-infection (c, n = 21, 2 replicates). This is also provided here for reviewers but has not been added to the paper.



We therefore propose to this reviewer that we have addressed the primary concern with biochemical experiments showing direct binding between Imp1 and a cargo protein (that is disrupted by mutants isolated from our screen).

Further comments:
The first sentence on page 4 „These proteins lack…" makes a logical connection that is not justified.
Modified to read (L67):
"Imp1-Imp5 have not been previously studied and have no obvious predicted molecular function. Imp6, an RNA-binding protein (Enustun et al. 2023), is additionally identified here as a direct physical interactor of Imp1. Together, we propose that these proteins constitute a unique phage protein trafficking system that enables transit into the phage nucleus. "

Page 5, line 83: "… a localization consistent with a role in protein import." Why actually? If Imp1 is active during protein import, I would rather expect a ring staining of the nucleus or an accumulation in the nucleus. Furthermore, the fact that the Imp1-mNeonGreen construct was inactive indicates that the results with this protein are questionable anyway.
This is a good point. We have removed the latter part of this offending statement to stay more focused on reporting what the localization of this protein is. As described above, we put most importance on the genetic results and new biochemical interactions reported but feel remiss if we were to not investigate its localization in any way. As mentioned in the paper, other studies have assessed localization of Imp1 homologs because of its conservation, where they have made similar conclusions about it appears, but have never assessed complementation.

Abbreviation mNG is not introduced.
Corrected

Information on the numerical aperture defining the optical resolution of the used microscope objective lens is missing.
The numerical aperture is: 1.450 and this was added to the methods section. (L617)

Ulrich Kubitscheck

Referees' comments:

Referee #1 (Remarks to the Author):

The authors have admirably addressed my concerns. The softened language and the new cross-plating experiment both constrain the claim to what the data actually show, and further support that claim. I also appreciated the additional discussion that explained why certain aspects were kept out of the manuscript.

I have one minor remaining point about the rates of escape, see below, and a few additional comments about clarifying the magnitude of some effects in text. The rest are simply curiosity and copy-edit issues, and the authors (and editors) can disregard as their leisure – just proving I read it!

Minor Comments:
L90 “substantially decreased by this mutation” – quantify/statistics?
Rephrased to remove substantially, Line 118.
L107 “very low titre, however” this isn’t a very strong statement. Do you mean that plaques had low pfu counts? Or that amplifications consistently yielded low titres? Not much depends on this statement, but not entirely clear what is being quantified here.
Clarified in the paper to summarize that both situations yielded low titer. Line 151.
L124: “robust phage targeting” magnitude?
Removed the phrase in question and replaced it with a real number. Line 168.
L131: “lower frequency of ~10^-7.” Given the frequency stated earlier was “10^6 to 10^7” it really is hard to know how much “lower” this frequency is.
Thank you for pointing this out. We were referring to multiple Nlp constructs with the 10^-6 to -7 statement. For this one (nlp1) specifically, this new isolation was indeed a lower frequency. Clarified in the text at line 222. We also added a new figure (Extended Data Fig 1a) with real plaque counts and EOPs for every targeting construct used in the paper.

Curiosity:
L72: Import and Impart. Clever. I like it.
L155: Double mutants no apparent fitness cost?
No obvious fitness defect was observed, but not deeply characterized for any mutant. To be consistent, we did not change the text here as we pointed out one other time when a mutant was specifically unfit (low titer in *imp2* nonsense mutation) but have not addressed this for any others.
L165: Do relevant residues co-evolve with e.g. some residues of the Nlps?

We agree that it would be interesting to know this, but seems we'd need a better understanding of binding surfaces on Nlps… we hope that our study will generate first steps in this direction for future work.
L184: The genetically dominant mutations are a neat phenotype. This isn't a question, just a statement.
Agree!

Copy Editing:
L56: This sentence is awkward, as it implies the Imp2/4/5 do not have a role in protein import. Consider rewording entirely. Changed
L76: Tense change (provides) a little awkward. Changed
L95: mNG already introued in line 91. Changed
L96: "and likely first initiates protein import" not clear where this is coming from (I don't disagree that import probably starts there, but there's neither a citation, nor experiment, nor reasoning provided")
An experiment has been added to Extended Data Fig 3, accompanied with a supplemental video (around Line 125), showing early Imp1-mNG localization at the site of infection at the cell pole.
L112: imp2 is underlined Removed
L115: "Imp2 only appears to be conserved" reword "imp2 appears to only be conserved"
Changed
L127: I think use of non-redundant didn't help here, consider rewording e.g. "none of which overlapped with those previously isolated"? Changed
L134: line break before section not present (was in prior section). Also missing before next section.
Fixed
L192: Why not mention the ratio of imp1 mutations to imp3 mutations, as for the earlier screen?
Changed to mention the numbers of mutants, Line 302.
References still incompletely edited, e.g. 6 lacks ital, 7 lacks ital and is in title case, 7 lacks ital and is title case, 11 lacks ital and is in title case, 12 in title case, 14 lacks capitalization of pseudomonas and ital, 17 lacks ital, 18 lacks ital, 19 lacks ital,
Fixed.

Referee #2 (Remarks to the Author):

The authors have nicely addressed my concerns. One additional suggestion is to add time stamps onto the live-cell movie.
A time stamp has been added to the live-cell movie.

Referee #3 (Remarks to the Author):

I appreciate your thorough responses to my questions. The clear emphasis on the specificity of the transport system into the phage nucleus as the main finding of the study highlights the significance of the meticulous work you have conducted. I am pleased to recommend the manuscript for publication.
Fantastic, thank you!

Please address the following two minor points before finalizing:

1.The paragraph lines 68-78 could be clearer. The current wording may be confusing as it first mentions four fused nuclear proteins and only later clarifies that these were selected from a larger pool. It would be beneficial to rephrase this section to explicitly state that a broader set of Nlps was initially tested, from which only four showed a significant reduction in phage titer. This adjustment would help clarify the sequence of events for the reader.
Clarified to match the reviewer's suggestion. Now starting at Line 77.

2. A minor point regarding Figure 2: Please add information about which phages were cultured in the figure description. I assume they were labeled by the colors corresponding to the bacterial species, but it would be helpful to make this explicit in the description.
To address this comment, we have added this sentence to the figure legend: "Cultured phages are colored by host species, and several representative phages are also labeled by name. Uncolored phages are from metagenomic sequencing." Line 682

Referee #4 (Remarks to the Author):

Here, I will only respond to the rebuttal letter and the additional experiments, which have been done in response to my original comments.

The direct evidence of binding between Imp1 and its cargo Nlp2 certainly validates the overall approach used in this work, which is further supported by the results with the various mutants mentioned.
Thank you for this comment.

However, in my first review I was not referring to the binding between Imp1 and its cargo, but to the binding between Imp1 and the components of the shell. Enustun et al. (2023) (ref 9) identified six proteins that associate with the nuclear shell, including gp2, gp61, gp63 and gp64. Demonstrating the direct interaction of imp1 with one (or more) of these proteins

would greatly strengthen the manuscript. More importantly, the identification of an interaction between a shell protein and imp1 would give a first indication of what kind of selection mechanism is present and how directionality is achieved in this intriguing system.
Ulrich Kubitscheck

Thank you for this clarification of what the reviewer was looking for.
We want to first respond to something the reviewer said: "Enustun et al. (2023) (ref 9) identified six proteins that associate with the nuclear shell, including gp2, gp61, gp63 and gp64. Demonstrating the direct interaction of imp1 with one (or more) of these proteins would greatly strengthen the manuscript." In fact, gp63 is Imp1 and gp61 is Imp6, in this phage. We apologize if this is not more clear in the paper and have tried to update the section around L125-146 to clarify this and unite it with the microscopy to suggest an interaction with the nuclear lattice. Therefore this means that proximity label interactions and pulldowns have connected Imp1/Imp6 and the nuclear shell protein (ChmA) already.

Regardless, in the interest of addressing the reviewer's comment we performed an *in vitro* pull-down of His-MBP-ChmA and Imp1. As ΦKZ's ChmA protein has never been purified or characterized *in vitro,* and other groups have reported difficulties with solubility when expressing the untagged ΦPA3 homolog (Nieweglowska et al., 2023, Enustun et al., 2023) an MBP tag was added to increase solubility in case issues with expression arose, as done in Nieweglowska et al. where they structurally characterized the ΦPA3 homolog. Encouragingly, His-MBP-ChmA appears to be very well expressed and soluble (Reviewer Figure 1).

His-MBP-ChmA (hereafter referred to as ChmA), Imp1-His, sfCherry-FLAG (negative control for ChmA non-specific binding) and His-MBP (negative control for Imp1 non-specific binding) were expressed in *E. coli* separately and lysed together in the appropriate pairings (please refer to Reviewer Figure 2). We performed MBP pull-downs. Elutions were analyzed by western blotting and were run through size exclusion chromatography, to assess the oligomeric state of eluted ChmA.

Two replicates showed Imp1 co-eluted with ChmA, but a fainter Imp1 band was also consistently detected in the MBP negative control elution. This suggests that Imp1 co-elutes with both ChmA and MBP, perhaps because it is non-specifically binding the MBP amylose resin at a low level (Reviewer Figures 2-4). This non-specific binding could not be removed by adding low concentrations of the eluent maltose to the wash (0.001mM, chosen from performing a maltose titration; manufacturer recommends eluting with 10-50mM maltose).

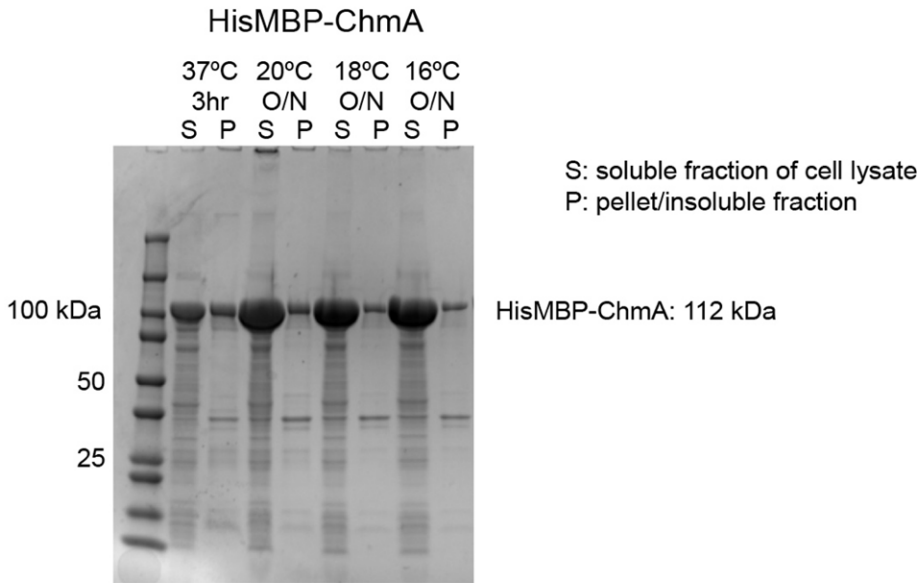
Although the non-specific binding limits our ability to make a confident conclusion at this time, we nonetheless ran the elution on SEC. We were encouraged to observe that Imp1 co-eluted with ChmA in a high molecular weight fraction, suggesting there may be a real interaction, as the Imp1 that non-specifically eluted with MBP came off the column much

later (not shown). However, in the end we have not put any conclusions about this experiment in the paper because of this non-specific binding.

In future work we will continue to optimize the pull-down conditions to answer this important question, perhaps changing protein tags, testing ChmA truncations, and attempting the same experiment with homologous proteins. Despite this being a very important and interesting question, we would argue that the evidence presented in our paper and referenced from previous work make a good case for an association with the nuclear lattice and we would respectfully request that we not delay publication of other exciting data in the manuscript.

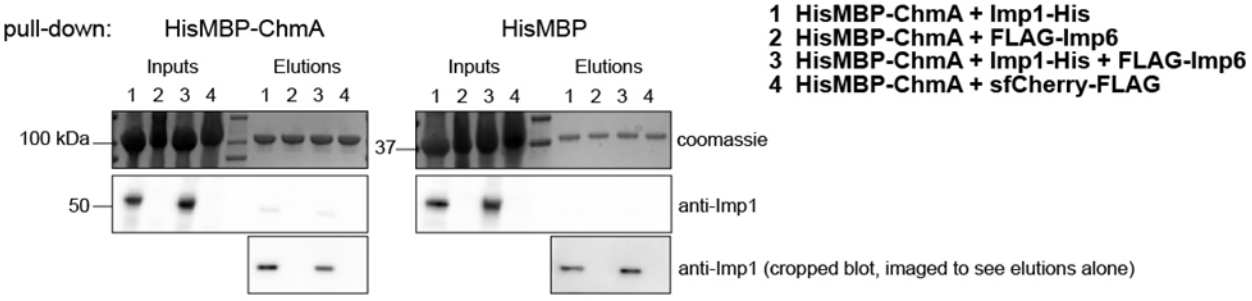
Reviewer Figure 1

HisMBP-ChmA is well-expressed



Reviewer Figure 2

Imp1 co-elutes with HisMBP-ChmA and HisMBP (-ctrl) in an MBP pull-down, suggesting non-specific binding



Summary: Imp1 is coming down more with MBP-ChmA alone vs. MBP but this non-specific binding is concerning and limits an exciting conclusion. Next slide has full Imp1 blot

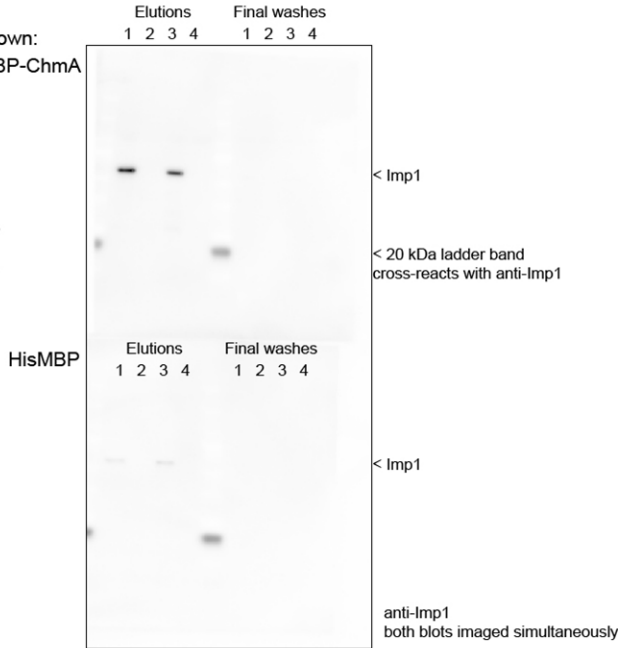
Reviewer Figure 3

Imp1 co-elutes with HisMBP-ChmA and HisMBP (-ctrl) in an MBP pull-down, suggesting non-specific binding

While Imp1 co-elutes with both HisMBP-ChmA and -ctrl HisMBP, when the two blots from the previous slide are imaged simultaneously, the relative amount of Imp1 pulled down with ChmA appears larger than with MBP alone.

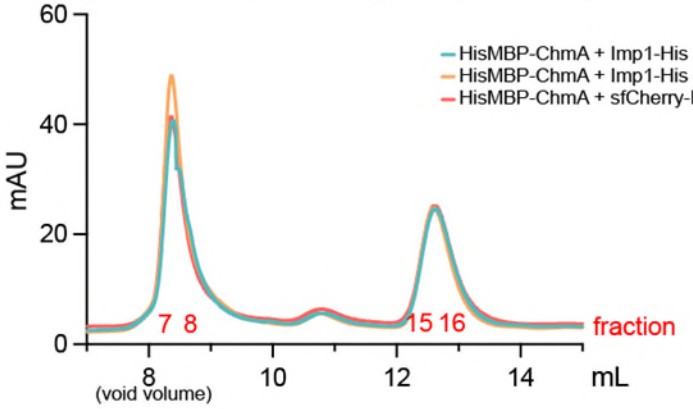
At present, we think these results are inconclusive and perhaps an alternate experimental setup and optimization is required to address this question.

Two additional final washes with 0.001mM maltose (25mM used to elute MBP tag from amylose resin) do not improve the pull-down specificity.



Reviewer Figure 4

Size exclusion chromatography shows Imp1 co-elutes with HisMBP-ChmA as a high molecular weight species (near column void volume)



Summary: the Imp1 protein that comes down with MBP-ChmA elutes near the void volume, potentially as an oligomer or soluble aggregate, which also limits conclusions.