1 **Conserved jumbo phage factors required for protein import into a phage**

2 **nucleus**

3

4 Claire Kokontis[1], Timothy A Klein[1], Sukrit Silas[1], Joseph Bondy-Denomy[1,2]*

5

6 [1]Department of Microbiology and Immunology, University of California, San

7 Francisco, San Francisco, CA, 94158, USA

8 [2]Quantitative Biosciences Institute, University of California, San Francisco, San

9 Francisco, CA 94158, USA

10

11 *Corresponding Author: joseph.bondy-denomy@ucsf.edu

12 **Abstract**

13 Bacteriophages use diverse mechanisms to evade anti-phage defenses systems.

14 ΦKZ-like jumbo phages assemble a proteinaceous nucleus-like compartment that

15 excludes antagonistic host nucleases, while internalizing DNA replication and

16 transcription machinery[1,2,3,4]. The phage factors required for protein import and the

17 mechanisms of selectivity remain unknown, however. Here, we uncover an import

18 system composed of proteins highly conserved across nucleus-forming phages,

19 together with additional cargo-specific contributors. Using a genetic selection that

20 forces the phage to decrease or abolish import of specific proteins, we determine

21 that the import of five different phage nuclear-localized proteins (Nlp) all require

22 distinct interfaces of the same factor, Imp1 (gp69). Imp1 forms discrete puncta in the

23 phage nuclear periphery likely in complex with a direct interactor Imp6 (gp67), a

24 conserved protein encoded nearby. The import of some proteins, including a host

25 topoisomerase (TopA), additionally require Imp3 (gp59), a factor required for proper

26 Imp1 function. Three additional phage proteins (Imp2, Imp4, Imp5) are also required

27 for the import of two queried nuclear cargos, perhaps acting as specific adaptors. We

28 therefore propose a core import system including Imp1, Imp3, and Imp6 with the

29 highly selective Imp1 protein licensing transport through a protein lattice.

Main

Segregation of cytoplasmic and organelle activity, with regulated movement of biomolecules between them, is a fundamental aspect of eukaryotic life. DNA-containing organelles, long thought to be exclusive to eukaryotes, have been recently discovered during bacteriophage infection in multiple genera of Gram-negative bacteria including *Pseudomonas, Escherichia, Serratia, Erwinia, Vibrio,* and *Salmonella*[1,2,4–8]. A phage-produced proteinaceous nucleus-like structure compartmentalizes DNA replication and transcription, while translation occurs in the cytoplasm[1]. This compartment is primarily made of a single protein called Chimallin or PhuN and excludes host nucleases, such as CRISPR-Cas and restriction enzymes[3], but imports various phage proteins and at least one host protein[1,9]. How this selectivity is achieved and the mechanism of movement through the protein-based lattice[5,10,11] that forms the phage nucleus remain unaddressed. Understanding how proteins move from the cytoplasm into the phage nucleus, analogous to what is achieved by secretion systems or eukaryotic nuclear import, may reveal new fundamental biological mechanisms.

Transport of cargo proteins through the eukaryotic lipid nuclear membrane is mediated by binding between linear sequence motifs and adaptor proteins called importins that shuttle them through a massive ~100 MDa nuclear pore[12]. In the ΦKZ-like jumbo phages, there is no known signal that licenses proteins for import, no known adaptors, and no confirmed nuclear shell constituents that function as an entry point for phage proteins. Given the abundance of uncharacterized genes in the large ΦKZ-like genomes (i.e. > 200 kb) and the general insolubility of the protein-based nucleus limiting basic interaction approaches[13], we use an unbiased genetic selection to identify the requirements for protein import. The screen implicates three

3

55   proteins (Imp1, Imp3, Imp6) that are broadly conserved in nucleus-forming phages

56   with a role in protein import and an additional three (Imp2, Imp4, Imp5) that appear

57   to be conserved only among related *Pseudomonas*-infecting nucleus-forming

58   phages. These proteins lack predicted molecular functions, and therefore form a

59   unique phage protein trafficking system that enables transit into the phage nucleus.

60

61   **Identification of a phage protein required for nuclear cargo import**

62   The ΦKZ phage nucleus excludes the restriction enzyme EcoRI (278 amino

63   acids, 31.5 kDa), but fusion of EcoRI to a nuclear-localized protein gp152

64   (UvsX/RecA homolog, hereafter referred to as Nlp1) allows EcoRI to enter the

65   nucleus and cleave phage DNA[3]. This suggests that import is licensed by recognition

66   of nuclear cargo, rather than acting by excluding antagonistic host proteins. To

67   identify phage factors that are required for proteins to move into the nucleus after

68   synthesis in the cytoplasm, we used a genetic selection experiment. We individually

69   fused four nuclear proteins, Nlp1 (gp152, RecA-like), Nlp2 (gp155, RNase H-like),

70   Nlp3 (gp104), and Nlp4 (gp171) to the C-terminus of EcoRI-sfCherry2 constructs

71   (hereafter referred to as EcoRI-NlpX fusions) to impart selective pressure on the

72   phage to reduce or abolish import of these cargo via mutations in putative import

73   factors[1,9] (Fig. 1a). All fusions reduced phage titer by ~10^6-fold due to the 92 EcoRI

74   recognition sites in the phage genome, while dead EcoRI (E111G) fused to Nlp1 also

75   localizes inside the phage nucleus but provides no reduction in titer[3]. EcoRI fusions

76   to other nuclear proteins (gp50, gp70, gp118, gp123, or gp179)[1,9] were inactive

77   (Extended Data Fig. 1a). An unmutagenized phage population with a titer of ~10^{11}

78   pfu/mL was subjected to targeting by each of the four EcoRI fusions and

79   spontaneous mutant escape phages emerged at a frequency of ~10^{-6}-10^{-7}. EcoRI-

80   Nlp2, EcoRI-Nlp3, and EcoRI-Nlp4 fusions all selected for distinct mutations in a

81   single uncharacterized phage gene, *orf69*, hereafter referred to as *imp1* for import

82   gene 1 (Fig. 1b, Table 1, Extended Data Table 1). A total of 8 distinct missense

83   mutant alleles in *imp1* were identified in this way, with no overlap (Table 1, Extended

84   Data Table 1). Complementation with WT *imp1* expressed *in trans* allowed EcoRI-

85   Nlp2/3/4 fusions to restrict *imp1* mutant escape phages again (Extended Data Fig.

86   1b), demonstrating that mutations in *imp1* are causal for escaping each EcoRI-Nlp

87   fusion.

88   To confirm that mutations in *imp1* decreased protein import, we imaged

89   EcoRI-Nlp2 during infection with WT phage or an *imp1 E310G* mutant. Indeed,

90   protein import was substantially decreased by this mutation (Fig. 1a, Extended Data

91   Fig. 2b-c). We next determined the subcellular localization of Imp1 with an

92   mNeonGreen fusion expressed from the host chromosome during infection with WT

93   ΦKZ. Upon infection, Imp1 went from being diffuse in the cell to forming 1-2 puncta

94   at the periphery of the phage nucleus (Fig. 1c, Extended Data Fig. 3a), a localization

95   consistent with a role in protein import. These tagged proteins, however, were not

96   able to resensitize imp1 mutant phages to EcoRI-Imp2 targeting (Extended Data Fig.

97   3b). This suggests that Imp1-mNG is inactive, perhaps because the fusion blocks

98   interactions with cargo or other binding partners. The ΦPA3 homolog of Imp1 (ΦPA3

99   gp63) was also recently identified as an interactor with Chimallin/PhuN through

100   proximity labeling in phage ΦPA3[9]. Microscopy experiments with the ΦPA3 Imp1

101   homolog as well the phage 201Φ2-1 Imp1 homolog (201-Φ2-1 gp125) similarly

102   revealed puncta in the nuclear periphery[1,9]. Together, these data suggest that Imp1

103   plays an important role in protein import into the phage nucleus.

104   Phages that escaped EcoRI-Nlp1 restriction, on the other hand, had 8 distinct

105   missense or nonsense mutations in a second uncharacterized gene, *orf47* (*imp2*)

106   (Fig. 1d, Table 1, Extended Data Table 1). EcoRI-Nlp1 import into the phage nucleus

107   was indeed disrupted during infection with an *imp2 K45N* mutant phage (Fig. 1a,

108   Extended Data Fig. 2d-e). *imp2* expression *in trans* also restored EcoRI-Nlp1 import

109   and targeting of the phage, confirming causality of this mutation (Extended Data Fig.

110   1c). Imp2 only appears to be conserved among related *Pseudomonas*-infecting

111   nucleus-forming phages based on sequence searches, and none of its homologs

112   have any known function (Extended Data Fig. 1d). The importance of Imp2 in Nlp1

113   import was confirmed by targeting a related phage, ΦPA3, with EcoRI-Nlp1, which

114   similarly revealed phage escape mutations in *orf44*, encoding an Imp2 homolog

115   (Extended Data Fig. 1e). Since Imp2 mutations were uniquely selected for by EcoRI-

116   Nlp1, we next wondered if Imp2 was also an imported protein. We fused Imp2 to

117   mNeonGreen to assess its localization during phage infection and found that it is

118   indeed imported into the nucleus (Fig. 1e). This suggests that Imp2 is required for

119   Nlp1 import but is itself imported. We therefore repeated the selection, fusing EcoRI

120   to Imp2, which resulted in robust phage targeting while dead EcoRI-Imp2 did not. We

121   found that phages that escaped EcoRI-Imp2 targeting contained new missense

122   mutations in *imp1* (Fig. 1f, Table 1, Extended Data Table 1). To understand why

123   mutations in *imp1* were not initially seen under EcoRI-Nlp1 selection, we expressed

124   Imp2 *in trans,* preventing successful phage escape via mutations in this gene.

125   Indeed, this approach selected directly for a mutation in *imp1* (*A270D*) (Fig. 1g). This

126   new class of Nlp1 import-deficient mutant phage emerged at a lower frequency of

127   $\sim 10^{-7}$, explaining why it wasn't initially seen. Together, these data demonstrate that

128   Nlp1 uniquely requires Imp2 for its import, while import of Nlp1-Nlp4 and Imp2 all

129   require Imp1 (Fig. 1h).

130

131   **Imp1 has distinct interfaces required for import of different phage nuclear**

132   **cargo**

133   The selective pressure to reduce import of Nlp1 and Imp2 selected for 15 unique

134   mutant alleles in *imp1*, none overlapping with the 8 *imp1* mutant alleles previously

135   isolated by the Nlp2, Nlp3, or Nlp4 fusions (Table 1, Extended Data Table 1). Given

136   the convergence on Imp1, we next queried whether the unique mutations decreased

137   import of all cargo or only some. Cells expressing each of the EcoRI-Nlp fusions

138   were infected with WT ΦKZ or various *imp1* mutant phages. Interestingly, individual

139   *imp1* mutations only decreased the phage targeting activity of the EcoRI fusion

140   construct they were selected under, suggesting specific perturbations to import (Fig.

141   2a). This suggests that the *imp1* mutation acquired by each escape phage likely

142   maintains sufficient expression, localization and import function for other cargo while

143   only specifically perturbing the import of one protein. To visualize the location of

144   functional residues, an AlphaFold2 predicted structural model of Imp1 was

145   generated, and each set of cargo-specific mutations labeled (Fig. 2b). The four

146   imported cargo selected for Imp1 mutations that generally clustered into five regions

147   on the Imp1 model. Consistent with discrete functional interfaces on Imp1, double

148   mutants could be isolated by iterative exposures to different EcoRI-Nlp fusions. For

149   example, an Imp2 import-deficient mutant phage (*imp1 G88D/R90H*) acquired a new

150   mutation in *imp1* (*R306L*) when selected on EcoRI-Nlp2 (Fig. 2c). The inverse was

151   also true, a phage that resisted import of EcoRI-Nlp2 (*imp1 H300Y*) acquired a

152   second mutation in *imp1* (*I226T*) when selected on EcoRI-Imp2. Together, we

7

153   propose that Imp1 is equipped with distinct interfaces to achieve import specificity

154   and licensing.

155   Sequence and structure-based searches with Imp1 did not reveal homologs

156   with any known molecular function. Imp1 is widely conserved in nucleus-forming

157   phages infecting diverse genera including *Pseudomonas, Serratia, Xanthomonas,*

158   *Aeromonas, Vibrio, Erwinia,* and members of the Enterobacteriaceae family (Fig.

159   2d). Nearly every phage genome encoding an Imp1 protein also encodes a homolog

160   of ΦKZ gp54/Chimallin/PhuN, the major phage nucleus protein. Only a few

161   exceptions are noted and likely due to incomplete genomic records. Together with

162   the genetic data demonstrating the importance of multiple Imp1 interfaces in the

163   import of distinct cargo, its localization during infection, and its broad conservation in

164   nucleus-forming phages, we propose that Imp1 is a key determinant of protein

165   transport into the phage nucleus for this phage family.

166

167   **Imp1 genetic and physical interactions with *imp3* and Imp6**

168   To identify factors required for Imp1 function or localization, we fused EcoRI

169   (or dead EcoRI) to Imp1, to assess its proximity to phage DNA and select for

170   mutants that disrupt its localization. Dead EcoRI-Imp1 did not reduce phage titer,

171   while active EcoRI-Imp1 abolished phage replication such that no escape mutants

172   could be isolated (limit of detection: 2 x 10^{-9}) (Extended Data Fig. 4a). This

173   demonstrates that the Imp1 N-terminus is localized inside the phage nucleus and

174   suggests that mutations that alter this localization may be lethal. Fusing EcoRI to 11

175   different Imp1 mutant proteins, with representatives chosen from selection

176   experiments with the various EcoRI-Nlp fusions, also maintained full phage targeting

177   with no escape mutants isolated (Extended Data Fig. 4a). This suggests that mutant

178   Imp1 proteins express and localize correctly, as proposed above.

179   To identify additional *imp1* genetic interactions not seen from the previous

180   selections, we used two alternative selection approaches, which both led to the

181   same new uncharacterized gene, *orf59* (*imp3*) (Fig. 3a-c). Using the EcoRI-Imp2

182   fusion that initially yielded *imp1* mutant phages, we repeated the selection but

183   expressed WT Imp1 *in trans* to enable isolation of phages with mutations in other

184   import factors. One class of mutant phages emerged that acquired spontaneous

185   missense mutations in *imp1* (Extended Data Fig. 3b, Table 1, Extended Data Table

186   1), which are genetically dominant. However, most phage mutants that escaped this

187   targeting construct acquired mutations in *imp3* (Fig. 3a). Mutations were either

188   coding changes or 10 bp upstream of the ATG start, in the likely Shine-Dalgarno

189   sequence (Fig 3c, Table 1, Extended Data Table 1). These data demonstrate that

190   *imp3* is required for correct Imp1 function with respect to Imp2 import. In a

191   complementary approach, we fused EcoRI to gp63 from phage ΦPA3, an Imp1

192   homolog with 52% a.a. identity (Imp1$_{ΦPA3}$). EcoRI-Imp1$_{ΦPA3}$ provided >100-fold

193   weaker selection pressure against ΦKZ than EcoRI-Imp1$_{ΦKZ}$ and enabled mutant

194   ΦKZ phages to emerge (Fig. 3b, Table 1, Extended Data Table 1). This approach

195   also selected for dominant *imp1* mutations and mutations in *imp3*, which possibly

196   perturb Imp1$_{ΦPA3}$ localization or assembly (Fig. 3c, Extended Data Fig. 3b).

197   Only one host protein is known to be localized inside the phage nucleus,

198   topoisomerase TopA[1]. The EcoRI-TopA fusion also selected for escape phages

199   (~10^{-6} frequency) with mutations in *imp1* and *imp3* (Fig. 3c-d, Extended Data Fig.

200   3b). Additionally, EcoRI-TopA selected for mutations in two new uncharacterized

201   genes, *orf48* (*imp4*) and *orf287* (*imp5*) (Fig. 3d). Like Imp2, identified above as a

202   protein specifically required for Nlp1 import, Imp4 and Imp5 are conserved only in

203   ΦKZ-like phages infecting *Pseudomonas* and no other EcoRI fusion selected for

204   mutations in these genes (Extended Data Fig. 5a). Thus, we interpret this to mean

205   that Imp4 and Imp5 are specifically required for TopA import, while core import

206   proteins Imp1 and Imp3 are additionally required.

207   To understand the contribution of Imp3 to protein import, we first assessed its

208   phylogenetic distribution and localization. Like Imp1, Imp3 is well-conserved among

209   nucleus-forming phages beyond those that infect *Pseudomonas*[6]. However,

210   fluorescently tagged Imp3 did not express well and thus its localization could not be

211   assessed. Additionally, EcoRI-Imp3 fusions were inactive (data not shown) and thus

212   whether it is proximal to phage DNA could not be ascertained. Finally, we suspect

213   that co-encoded genes in a putative operon[14] with *imp3* are required for proper

214   expression or function as these genes (*p18-imp3-orf60-orf61*) were required to

215   achieve partial complementation EcoRI-Imp1$_{\Phi PA3}$ and EcoRI-TopA escape

216   (Extended Data Fig. 4c, 5b). The functional contributions of *imp3* and its gene

217   neighbors will require further investigation.

218   Finally, we reasoned that the process of protein import may require additional

219   proteins that associate with Imp1 that were not identified by genetic selection

220   experiments. To find candidate interactors, we examined the operon structure and

221   conservation of the locus around *imp1* (*orf69*). This locus encodes gp67-70 (Fig. 1b),

222   but previous work has established that gp68 (together with gp71-73 and gp74) is part

223   of the nvRNAP complex[15]. Therefore, we assessed whether Imp1 interacts with gp67

224   or gp70. 6xHis-tagged Imp1 expressed in *E. coli* immunoprecipitated with gp67-

225   FLAG, but not gp70-FLAG (Fig. 4a) and a reciprocal FLAG immunoprecipitation

226   revealed the same result (Fig. 4b). To confirm that Imp6-FLAG doesn't adhere to Ni-

227   NTA beads non-specifically, 6xHis-tagged Nlp1 was used, which did not pull-down

228   gp67-FLAG (Fig. 4a). The Imp1-gp67 interaction was stable over size exclusion

229   chromatography (Fig. 4c). Its elution position and mass photometry suggested a

230   ~120 kDa complex consistent with a stoichiometry of $gp67_2$:$Imp1_1$ (Fig. 4d). We

231   therefore propose naming gp67 as Imp6, a direct binder of the core Imp1 protein and

232   a putative member of a phage nuclear import complex.

233

234   **Discussion**

235   The ΦKZ-like family of jumbo phages assembles a remarkable protein-based

236   nucleus during infections, which selectively internalizes or excludes proteins,

237   demonstrating a eukaryotic nucleus-like segregation ability. These phages often

238   have very large genomes (> 200 kb) with many genes of unknown function that likely

239   contribute to this and other processes executed by this phage family, necessitating

240   unbiased approaches to identify gene functions. Through the deployment of a

241   genetic selection approach, we have identified 5 factors (Imp1-5) required for the

242   import of 6 proteins (Nlp1-4, Imp2, TopA).

243   A key finding from this work is the functional identification of an import factor

244   Imp1, a previously uncharacterized protein that is broadly conserved in nucleus-

245   forming jumbo phages. Imp1 is a key specificity determinant for licensing import, as

246   demonstrated by the isolation of individual mutations that specifically abolish the

247   import of distinct cargo proteins. However, it currently remains unclear how proteins

248   physically transit into the phage nucleus. Possible options include a bona fide pore, a

249   flippase-like mechanism, or an unfoldase that participates in threading the cargo

250   through the protein lattice. We additionally report a functional role for a protein

251   encoded nearby, Imp3, and a direct physical interaction between Imp1 and Imp6. An

252   Imp6 homolog in phage ΦPA3 was recently described as a non-specific RNA binding

253   protein localized to the periphery of the phage nucleus[16]. Future work is needed to

254   determine the connection between non-specific RNA binding and potential

255   contributions to RNA export/localization and protein import.

256   The fundamental biological challenge of protein localization in segregated

257   organellar compartments has largely been studied in eukaryotic systems. This phage

258   family and its protein-based nucleus provide a new challenge and opportunity for

259   understanding the basic mechanisms of protein movement. The identification of

260   phage genes required for import, along with a method for their discovery in other

261   phage systems, are key first steps to unraveling this fascinating mystery.

References

262

263   1. Chaikeeratisak, V. *et al.* Assembly of a nucleus-like structure during viral

264   replication in bacteria. *Science* **355**, 194–197 (2017).

265   2. Chaikeeratisak, V. *et al.* The Phage Nucleus and Tubulin Spindle Are Conserved

266   among Large Pseudomonas Phages. *Cell Rep.* **20**, 1563–1571 (2017).

267   3. Mendoza, S. D. *et al.* A bacteriophage nucleus-like compartment shields DNA

268   from CRISPR nucleases. *Nature* **577**, 244–248 (2020).

269   4. Malone, L. M. *et al.* A jumbo phage that forms a nucleus-like structure evades

270   CRISPR--Cas DNA targeting but is vulnerable to type III RNA-based immunity.

271   *Nature microbiology* **5**, 48–55 (2020).

272   5. Laughlin, T. G. *et al.* Architecture and self-assembly of the jumbo bacteriophage

273   nuclear shell. *Nature* **608**, 429–435 (2022).

274   6. Prichard, A. *et al.* Identifying the core genome of the nucleus-forming

275   bacteriophage family and characterization of Erwinia phage RAY. *Cell Rep.* **42**,

276   112432 (2023).

277   7. Weintraub, S. T. *et al.* Global Proteomic Profiling of Salmonella Infection by a

278   Giant Phage. *J. Virol.* **93**, (2019).

279   8. Jacquemot, L. *et al.* Therapeutic Potential of a New Jumbo Phage That Infects

280   Vibrio coralliilyticus, a Widespread Coral Pathogen. *Front. Microbiol.* **9**, 2501

281   (2018).

282   9. Enustun, E. *et al.* Identification of the bacteriophage nucleus protein interaction

283   network. *Nat. Struct. Mol. Biol.* **30**, 1653–1662 (2023).

284   10. Nieweglowska, E. S. *et al.* The ΦPA3 phage nucleus is enclosed by a self-

285   assembling 2D crystalline lattice. *Nat. Commun.* **14**, 927 (2023).

286   11. Heymann, J. B. *et al.* The Mottled Capsid of the Salmonella Giant Phage

287   SPN3US, a Likely Maturation Intermediate with a Novel Internal Shell. *Viruses*

288   **12**, (2020).

289   12. Lin, D. H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update).

290   *Annu. Rev. Biochem.* **88**, 725–783 (2019).

291   13. Fossati, A. *et al.* Next-generation proteomics for quantitative Jumbophage-

292   bacteria interaction mapping. *Nat. Commun.* **14**, 5156 (2023).

293   14. Putzeys, L. *et al.* Refining the transcriptional landscapes for distinct clades of

294   virulent phages infecting pseudomonas aeruginosa. *bioRxiv* 2023.10.13.562163

295   (2023) doi:10.1101/2023.10.13.562163.

296   15. Yakunina, M. *et al.* A non-canonical multisubunit RNA polymerase encoded by a

297   giant bacteriophage. *Nucleic Acids Res.* **43**, 10411–10420 (2015).

298   16. Enustun, E. *et al.* A phage nucleus-associated RNA-binding protein is required

299   for jumbo phage infection. *bioRxiv* (2023) doi:10.1101/2023.09.22.559000.

300   17. Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P. & Yu, H. D. PBAD-based

301   shuttle vectors for functional analysis of toxic and highly regulated genes in

302   Pseudomonas and Burkholderia spp. and other bacteria. *Appl. Environ.*

303   *Microbiol.* **74**, 7422–7426 (2008).

304   18. Choi, K.-H. & Schweizer, H. P. mini-Tn7 insertion in bacteria with single attTn7

305   sites: example Pseudomonas aeruginosa. *Nat. Protoc.* **1**, 153–161 (2006).

306   19. Choi, K.-H. *et al.* Genetic tools for select-agent-compliant manipulation of

307   Burkholderia pseudomallei. *Appl. Environ. Microbiol.* **74**, 1064–1075 (2008).

308   20. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat.*

309   *Methods* **9**, 357–359 (2012).

310   21. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis.

311   *Nat. Methods* **9**, 676–682 (2012).

312 **Figure 1: Isolation of mutant phages that reduce protein import into phage**

313 **nucleus**

314 **a**, Live-cell fluorescence microscopy of *P. aeruginosa* strain PAO1 expressing the

315 indicated sfCherry2 fusion proteins infected with WT ΦKZ or the indicated phage

316 mutant. DAPI stain indicates phage DNA within the phage nucleus (white arrow). **b**,

317 **d**, **f**, Plaque assays with the indicated WT or mutant phage spotted in 10-fold serial

318 dilutions on a lawn of PAO1 expressing indicated sfCherry2 fusion. dEcoRI,

319 catalytically inactive EcoRI. **c**, **e**, Live-cell fluorescence microscopy of Imp1 fused to

320 mNeonGreen (mNG) infected with WT ΦKZ (n = 94) (**c**) or Imp2 fused to

321 mNeonGreen and infected with WT ΦKZ (n = 129) (**e**). **g**, Plaque assays with the

322 indicated WT or mutant phage spotted in 10-fold serial dilutions on a lawn of PAO1

323 expressing the indicated sfCherry2 fusion, with or without expression of Imp2 *in trans*

324 (+/− Imp2). **h**, Schematic representing the factors required for nuclear import of

325 proteins queried in the EcoRI selections, as determined by the escape mutations

326 isolated in *imp1* and *imp2* from each selection (i.e. Nlp2 → Imp1 indicates Nlp2

327 requires WT Imp1 for its nuclear import). All plaque assays were replicated three or

328 more times. All microscopy experiments were replicated two or more times. Scale

329 bars, 1μm.

330

331 **Figure 2: Imp1 possesses distinct interfaces to enable import specificity**

332 **a**, **c**, Plaque assays with indicated WT or *imp1* mutant phages spotted in 10-fold

333 serial dilutions on a lawn of PAO1 expressing the indicated sfCherry2 fusion.

334 dEcoRI, catalytically inactive EcoRI. Plaque assays were conducted as in Fig 1b. **b**,

335 AlphaFold2 high confidence predicted structural model of Imp1, with mutated

336 residues from isolated *imp1* mutant phages color coded by the EcoRI selection from

15

337 which they were isolated. **d**, Phylogenetic tree of Imp1 homologs. A closed black

338 circle indicates that no phage nuclear shell homolog was detected in the

339 corresponding phage genome by 3 iterations of PSI-BLAST or local blastp. Open

340 black circles indicate that the associated genome record is incomplete, and no

341 phage nuclear shell homolog was detected by the same methods. A closed circle

342 indicates that no nuclear shell homolog was detected by the same methods, but the

343 tubulin homolog responsible for centering the phage nucleus in the cell, PhuZ, was

344 detected by blastp. All plaque assays were replicated three or more times.

345

346 **Figure 3: Imp3 is required for proper Imp1 function**

347 **a**, **b**, Plaque assays with the indicated WT or mutant phage spotted on PAO1

348 expressing the indicated sfCherry2 fusion, with or without Imp1 expressed *in trans*

349 (+/– Imp1) (**a**); Imp1$_{\Phi PA3}$ indicates the phage ΦPA3 Imp1 homolog (**b**); dEcoRI,

350 catalytically inactive EcoRI. Plaque assays were performed as in Fig 1b. **c,**

351 AlphaFold2 high confidence predicted structural model of Imp3 with mutated DNA

352 bases or residues from isolated *imp3* mutant phages color coded by the EcoRI

353 selection from which they were isolated. **d**, Plaque assays with the indicated WT or

354 mutant phage spotted on PAO1 expressing the indicated sfCherry2 fusion. **e**,

355 Schematic representing the factors required for nuclear import of all proteins queried

356 in the EcoRI selections in this work, as determined by the escape mutations isolated

357 in *imp1-imp5* from each selection. All plaque assays were replicated three or more

358 times.

359

360 **Figure 4: Imp6 stably binds to Imp1**

16

361   **a**, **b**, Western blot analysis (anti-Imp1) of a His pull-down of His-tagged Imp1 with

362   FLAG-tagged Imp6, gp70, or Nlp1-His with Imp6-FLAG as a negative control (**a**) or a

363   FLAG pull-down of Imp1-His with FLAG-Imp6 or FLAG-gp70 (right) (**b**); I, input; E,

364   elution. **c**, Size exclusion chromatography (SEC) of Imp1-His alone or in complex

365   with Imp6-FLAG. Peak Imp1-His/Imp6-FLAG SEC fractions (indicated by mL

366   collected off the column) were run on an SDS-PAGE gel and stained with

367   coomassie. **d**, Estimation of Imp1-His/Imp6-FLAG complex molecular weight by

368   mass photometry.

369

370   **Methods**

371   **Bacterial growth and genetic manipulation**

372   *E. coli* strains XL1Blue and BL21 (DE3) and *P. aeruginosa* strain PAO1 were grown

373   in Lysogeny Broth (LB) medium at 37 °C shaking at 175 rpm. Bacteria were plated

374   on LB solid agar and any necessary antibiotics to maintain plasmids, with 10 mM

375   $MgSO_4$ when plating for phage infection. Expression of genes inserted in the

376   chromosomal attTn7 site was induced with 1 mM IPTG in the LB solid agar. Basal

377   expression of EcoRI-sfCherry2-NlpX or -ImpX fusions cloned into the pHERD30T[17]

378   plasmid under the leaky arabinose-inducible promoter was sufficient to induce phage

379   restriction and enable visualization by microscopy, and thus L-arabinose was not

380   added to plates or liquid cultures.

381

382   **Phage growth**

383   Phages were grown at 30 °C on LB solid agar with 10 mM $MgSO_4$ with any

384   necessary antibiotics and inducer. 150 µL of bacteria and 10 µL of phage were

385   mixed in 3 mL of 0.35% top agar with 10 mM $MgSO_4$ and plated on LB solid agar.

386   Plates were incubated at 30 ºC overnight. The next day, individual plaques were

387   picked and stored in 200 µL of SM phage buffer. Escaper phages were plaque

388   purified three times by repeating this method. High titer lysates were generated by

389   infecting PAO1 expressing the fusion construct that was used in selection overnight

390   in liquid LB 10 mM $MgSO_4$ with appropriate antibiotics and inducers at 37 ^{o}C. The

391   supernatant was collected and treated with 5% volume chloroform, shaken gently for

392   2 min, and spun down for 5 min at max speed to remove cell debris. This was

393   repeated once more, and the final phage lysate was stored with 1-5% volume

394   chloroform.

395

396   **Phage spot titration plaque assays**

397   150 µL PAO1 culture was mixed with 0.35% top agar and poured on solid LB agar

398   plates. Once solidified, 10-fold dilutions of phage in SM phage buffer were spotted

399   on the surface in 3 µL spots.

400

401   **Construction of fusion proteins**

402   The shuttle vector pHERD30T[17] was used for cloning and expression of EcoRI-

403   sfCherry2-Nlp/Imp and mNeonGreen fusions in *P. aeruginosa* strain PAO1. The

404   vector has a gentamicin selectable marker and an arabinose-inducible promoter. The

405   vector was digested with NcoI and HindIII restriction enzymes and purified. Inserts

406   were amplified by PCR from diluted ΦKZ lysates or PAO1 colonies as the DNA

407   template, and joined into the linearized vector by Hi-Fi Gibson DNA Assembly (NEB)

408   according to the manufacturer's protocol. The resulting reactions were used to

409   transform *E. coli* XL1-Blue competent cells. All plasmid constructs were sequenced

410   using primers that annealed to the vector outside of the multiple cloning site

411 (Quintara), or were whole-plasmid sequenced (Primordium). Plasmids were then

412 electroporated into competent PAO1 cells and selected on gentamicin.

413

414 **Construction of chromosomal insertion strains**

415 For chromosomal insertion of phage import factor genes into PAO1, genes of

416 interest were cloned into the pUC18-miniTn7T-LAC[18], once linearized with SpeI and

417 SacI and purified. Inserts were amplified by PCR from diluted ΦKZ lysates and

418 joined into the linearized vector by Hi-Fi DNA Assembly (NEB). The resulting vectors

419 were used to transform *E. coli* XL1-Blues and verified by sequencing using primers

420 that anneal to the vector outside of the multiple cloning site. The transposase helper

421 vector pTNS3[19] was used with miniTn7 constructs to transform PAO1 and insert the

422 genes of interest into the PAO1 chromosome at the Tn7 locus, with a –pTNS3

423 control in parallel, and the transformation was selected on gentamicin. Candidate

424 integrants were screened by colony PCR using PTn7R and PglmS-down, as well as

425 an internal gene-specific primer paired with PglmS-up. Integrants were then

426 amplified and sequenced with primers that annealed outside the attTn7 site to verify

427 integration. Electrocompetent cell preparation, electroporation, integration, FLP-

428 recombinase-mediated gentamicin marker excision using pFLP2 and plasmid curing

429 via sucrose counterselection were performed as described previously[18].

430

431 **Whole genome sequencing**

432 Genomic DNA extraction from phage lysates was performed by adding 200 µL of

433 lysis buffer (final concentration 10 mM Tris, 10 mM EDTA, 100 µg/mL Proteinase K,

434 100 µg/mL RNaseA, 0.5% SDS) to 200 µL high titer phage lysate (>10^9 PFU/mL)

435 and incubated at 37 ºC for 30 min, and then 55º C for 30min. Preps were then

19

436   purified by a phenol chloroform extraction followed by a chloroform extraction and

437   then ethanol precipitated, or by using the DNA Clean & Concentrator Kit (Zymo

438   Research). DNA was then quantified using the Qubit 4.0 Fluorometer (Life

439   Technologies). 50-100 ng genomic DNA was used to prepare WGS libraries using

440   the Illumina DNA Prep Kit (previously, Nextera Flex Library Prep Kit). A modified

441   protocol was used, with 5x reduced quantities of tagmentation reagents per prep,

442   except for the bead washing step where the recommended 100 μL of Tagment Wash

443   Buffer (TWB) was used. On-bead PCR indexing-amplification was performed using

444   custom-ordered indexing primers (IDT) matching the Illumina Nextera Index Kit

445   sequences and 2x Phusion Master Mix (NEB). PCR reactions were amplified for 9-

446   11 cycles, and subsequently resolved by agarose gel electrophoresis. DNA products

447   were excised around the 400 bp size range and purified using the Zymo Gel DNA

448   Recovery kit (Zymo Research). Libraries were quantified by Qubit. Libraries were

449   pooled in equimolar ratios and sequenced on the Illumina MiSeq using 150-cycle v3

450   reagents (Single end: 150 cycles, Read 1; 8 cycles, Index 1; 8 cycles, Index 2.

451   Paired end: 75 cycles, Read 1; 8 cycles, Index 1; 8 cycles, Index 2, 75 cycles, Read

452   2). Data were demultiplexed on-instrument, and trimmed using cutadapt (v1.15) to

453   remove Nextera adapters. Trimmed reads were mapped using Bowtie 2.0[20]. (–very-

454   sensitive-local alignments) and alignments were visualized using IGV (v2.11.0).

455   Mutations were called if they were present in >90% sequencing reads at loci with at

456   least 20x coverage.

457

458   **Live-cell fluorescence microscopy**

459   0.8% LB agar pads (25% LB, 2.5 mM $MgSO_4$) were supplemented with 0.5 μg/mL

460   DAPI for phage DNA staining. PAO1 strains expressing each of the fluorescent

461 protein constructs were grown in liquid culture supplemented with inducer if

462 necessary (0.05% arabinose for pHERD30T-mNG-Imp2, 0.1 mM IPTG for

463 Tn7::Imp1-mNG) to induce construct expression until OD600 ~0.5, and subsequently

464 infected with ΦKZ lysate for 50 min at 30 °C before imaging. Microscopy was

465 performed on an inverted epifluorescence (Ti2-E, Nikon, Tokyo, Japan) with the

466 Perfect Focus System (PFS) and a Photometrics Prime 95B 25 mm camera. Images

467 were acquired using Nikon Elements AR software (version 5.02.00). Cells were

468 imaged through channels of blue (DAPI, 50 ms exposure, for phage DNA), green

469 (FITC, 200 ms exposure, for mNeonGreen constructs), red (Cherry, 200 ms

470 exposure, for sfCherry2 constructs) and phase contrast (200 ms exposure, for cell

471 recognition) at 100x objective magnification. Final figure images were prepared in Fiji

472 (version 2.1.0/1.53c)[21].

473

474 **Phylogenetic analysis**

475 Homologs of ΦKZ import factors were identified by 3 iterations of PSI-BLAST against

476 the non-redundant protein database. Hits with >70% coverage and E-value <0.005

477 were included to generate a multiple sequence alignment (MSA) using MAFFT

478 (v7.490, fast strategy). The genomes not labeled as bacteriophages with homologs

479 were manually inspected for the presence of phage structural genes (i.e. to exclude

480 bacterial contigs), and were excluded from the MSA input if no phage genes were

481 annotated. After manual inspection of the MSA, results were input to FastTree

482 (v2.1.11 SSE3, default settings) to generate a phylogenetic tree and visualized in

483 iTOL: Interactive Tree of Life. To determine whether phage genomes containing

484 Imp1 homologs also encode Chimallin/PhuN (gp54 in ΦKZ), gp54 homologs were

485 acquired by 3 iterations of PSI-BLAST as described above, and genomes from Imp1

486   homologs were inspected for the presence of gp54 homologs. Genomes lacking an

487   apparent gp54 homolog according to PSI-BLAST were inspected manually to identify

488   a predicted gp54 locus (flanked by two well-conserved phage-encoded genes: DNA

489   polymerase upstream and RNA polymerase beta subunit downstream). These

490   candidates were confirmed as gp54 homologs by their similarity (>50% identity,

491   >70% coverage) to a homolog present in a genome on the Imp1 tree.

492

493   **Protein expression and purification**

494   6xHis-tagged Imp1 and Nlp1 were cloned into pET29b while FLAG-tagged Imp6 and

495   gp70 were cloned into pETduet-1. Binary interactions were assessed by co-

496   expression in BL21 (DE3) cells and purification by affinity chromatography using Ni-

497   NTA. For expression, 100 mL cultures of cells were grown in LB supplemented with

498   Kanamycin (50 µg/mL) and Ampicillin (100 µg/mL) at 37 °C shaking at 175 rpm.

499   When the cultures reached OD600 ~0.6, protein expression was induced with 1 mM

500   IPTG and incubated at 18 °C for ~16 hours. Cells were centrifuged at 5,000 x g and

501   resuspended in lysis buffer (25 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM

502   imidazole, 0.5 mM TCEP, protease inhibitor). Lysis was performed by sonication at

503   20% amplitude for 10 s, four times. Insoluble material was removed from the lysate

504   by centrifugation at 21,000 x g for 30 min. 300 µL of Ni-NTA resin slurry (Qiagen)

505   was washed with 10 mL of wash buffer (25 mM Tris-HCl pH 8.0, 300 mM NaCl, 10

506   mM imidazole, 0.5 mM TCEP) in a gravity column. Cleared lysate was then run over

507   Ni-NTA resin and non-specific interactors were removed by 10 mL washes using

508   wash buffer (four washes total). Protein was eluted with 300 µL of elution buffer (25

509   mM Tris-HCl pH 8.0, 300 mM NaCl, 400 mM imidazole). Pulldown experiments on

510   FLAG-tagged proteins were done similarly to those of His-tagged proteins but with

511 several modifications. Cells were resuspended in lysis/wash buffer (25 mM Tris-HCl,

512 pH 8.0, 300 mM NaCl, protease inhibitor) and sonicated as described above. 25 µL

513 of anti-FLAG magnetic agarose resin (Pierce) was washed and then incubated with

514 cleared lysate for one hour at 4 ⁰C with constant agitation. Protein bound resin was

515 washed five times with 1 mL of wash buffer using a magnetic separation rack. Bound

516 protein was eluted with 100 µL of 1.5 mg/mL FLAG peptide (Millipore) resuspended

517 in wash buffer. Protein interactions were assessed by Western blot (see below). Size

518 exclusion chromatography was performed with a Superdex 200 increase 10/300 GL

519 column using an AKTA Pure Protein Purification System (Cytiva). SEC buffer (25

520 mM Tris-HCl pH 8.0, 300 mM NaCl) was used for purification and protein samples

521 collected were subject to Western blotting and mass photometry.

522

523 **Western blotting**

524 The protein samples collected through pulldown and SEC analyses were mixed 3:1

525 with 4x Laemmli buffer supplemented with β-mercaptoethanol and boiled for 10

526 minutes. These samples were run on precast SDS-PAGE gels (Bio-Rad) and

527 transferred to PVDF membranes. The membranes were blocked with TBS-T buffer

528 (1x TBS, 0.1% Tween-20) supplemented with 5% skim milk. A custom primary

529 antibody for Imp1 (α-Imp1, Genscript), a commercial primary FLAG antibody (α-

530 FLAG, Millipore cat# F1804), or a commercial primary His antibody (α-His, Cell

531 Signaling cat# 2365S) were then added to the skim milk buffer at a titer of 1:5000

532 and left to incubate at room temperature for one hour. The membranes were washed

533 3x with 10 mL of TBS-T, then incubated with α-Rabbit (Cell Signaling, cat# 7074S) or

534 α-Mouse (Invitrogen, cat# 62-6520) secondary antibody at a titer of 1:5000 for 45

535 minutes at room temperature. The membrane was washed 3x with 10 mL of TBS-T

536   and developed with Clarity Max ECL substrate (Bio-Rad). Western blot images were

537   taken with an Azure Biosystems C400 imager.

538

539   **Mass photometry**

540   Imp1-Imp6 protein complexes purified by size exclusion chromatography were

541   analyzed by using a OneMP mass photometer (Refeyn). Adequate data collection

542   was made by mixing 1 µL of 1 µM protein with 15 µL of buffer (25 mM Tris-HCl, 300

543   mM NaCl). Data collection was done with the AcquireMP software (Refeyn 2024

544   R1.1). Data was collected for one minute and yielded 3055 measurable events. Data

545   processing was done using the DiscoverMP software (Refeyn DiscoverMP 2024 R1).

546

Materials and correspondence

Supplementary Information is available for this paper. Correspondence and requests

for materials should be addressed to Joseph Bondy-Denomy.

Author contributions

C.K. and J.B.-D. conceived the project and designed experiments. C.K. executed

genetic screens, sequencing, genetics, microscopy, and associated data analysis.

T.K. executed biochemical experiments. S.S. provided sequencing execution and

analysis. J.B.-D. and C.K. wrote the manuscript and all authors edited. J.B.-D.

supervised experiments and procured funding.

Declaration of interests

J.B.-D. is a scientific advisory board member of SNIPR Biome and Excision

Biotherapeutics, a consultant to LeapFrog Bio, and a scientific advisory board

member and co-founder of Acrigen Biosciences. The Bondy-Denomy lab received

research support from Felix Biotechnology.

572 **Extended Data Fig. 1**

573 **a-c**, Plaque assays with WT ΦKZ or WT 14-1, an EcoRI-sensitive phage (**a**) or

574 indicated WT ΦKZ and mutant phage (**b**, **c**) spotted in 10-fold serial dilutions on a

575 lawn of PAO1 expressing indicated sfCherry2 fusions or an empty vector (EV)

576 control, with or without expression of the appropriate phage gene *in trans* from the

577 bacterial attTn7 site (Tn7::*impX*) (**b, c**). **d**, Imp2 phylogenetic tree. **e**, Plaque assays

578 with the indicated ΦPA3 WT or mutant phage spotted in 10-fold serial dilutions on a

579 lawn of PAO1 expressing indicated sfCherry2 fusions. Plaque assays were

580 performed as in Fig 1b and replicated two times or more.

581

582 **Extended Data Fig. 2**

583 **a-e**, Representative images of live-cell fluorescence microscopy of PAO1 expressing

584 the indicated sfCherry2 fusions, infected with WT or indicated mutant ΦKZ. Scale

585 bars, 1 µm. Microscopy was performed as in Fig. 1c and replicated two times or

586 more.

587

588 **Extended Data Fig. 3**

589 **a**, Live-cell fluorescence microscopy of PAO1 expressing Imp1-mNeonGreen (Imp1-

590 mNG) from the attTn7 site, infected with WT ΦKZ (left, middle column; scale bar, 1

591 µm) or uninfected (right column; scale bar, 2 µm). Microscopy was performed as in

592 Fig 1c, and replicated three times. **b**, Plaque assays with the indicated WT or mutant

593 phage spotted in 10-fold serial dilutions on a lawn of PAO1 expressing the indicated

594 sfCherry2 fusions, with or without expression of the appropriate phage gene *in trans*

595 from the bacterial attTn7 site (Tn7::ImpX). Plaque assays were performed as in Fig

596 1b and replicated two times or more.

597

598 **Extended Data Fig. 4**

599 **a**, Plaque assays with WT ΦKZ or WT 14-1 (EcoRI-sensitive phage) spotted in 10-

600 fold serial dilutions on a lawn of PAO1 expressing the indicated sfCherry2 fusions. **b**,

601 High confidence AlphaFold2 predicted structural model with mutated residues from

602 isolated *imp1* mutant phages color coded by the EcoRI selection from which they

603 were isolated. **c**, Plaque assays with the indicated WT or mutant phage spotted on

604 PAO1 expressing the indicated sfCherry2 fusions, with or without Imp3 or the Imp3

605 operon (*p18-imp3-orf60-orf61*) expressed *in trans*. Plaque assays were performed as

606 in Fig 1b and replicated two times or more.

607

608 **Extended Data Fig. 5**

609 **a**, Imp4 and Imp5 phylogenetic trees. **b**, Plaque assays with the indicated WT or

610 mutant phage spotted in 10-fold serial dilutions on a lawn of PAO1 expressing the

611 indicated sfCherry2 fusions, with or without Imp3 or the Imp3 operon (*p18-imp3-*

612 *orf60-orf61*) expressed *in trans*. Plaque assays were performed as in Fig 1b and

613 replicated three times.

Table 1: Summary of mutant alleles selected for by EcoRI fusions

Selection	Mutated gene	Accession no.	No. of unique alleles
EcoRI-Nlp1	*imp2* (*orf47*)	NP_803613.1	8
Tn7::Imp2 [EcoRI-Nlp1]	*imp1* (*orf69*)	NP_803635.1	1
EcoRI-Nlp2	*imp1*		3
EcoRI-Nlp3	*imp1*		2
EcoRI-Nlp4	*imp1*		3
EcoRI-Imp2	*imp1*		14
	imp3 (*orf59*)	NP_803625.1	1
EcoRI-Nlp2+ EcoRI-Imp2*	*imp1*		10
Tn7::Imp1 [EcoRI-Imp2]	*imp1*		2
	imp3		9
EcoRI-Imp1ΦPA3	*imp1*		2
	imp3		4
EcoRI-TopA	*imp1*		2
	imp3		2
	imp4 (*orf48*)	NP_803614.1	4
	imp5 (*orf287*)	NP_803853.1	1

* indicates selections were performed sequentially

Figure 1
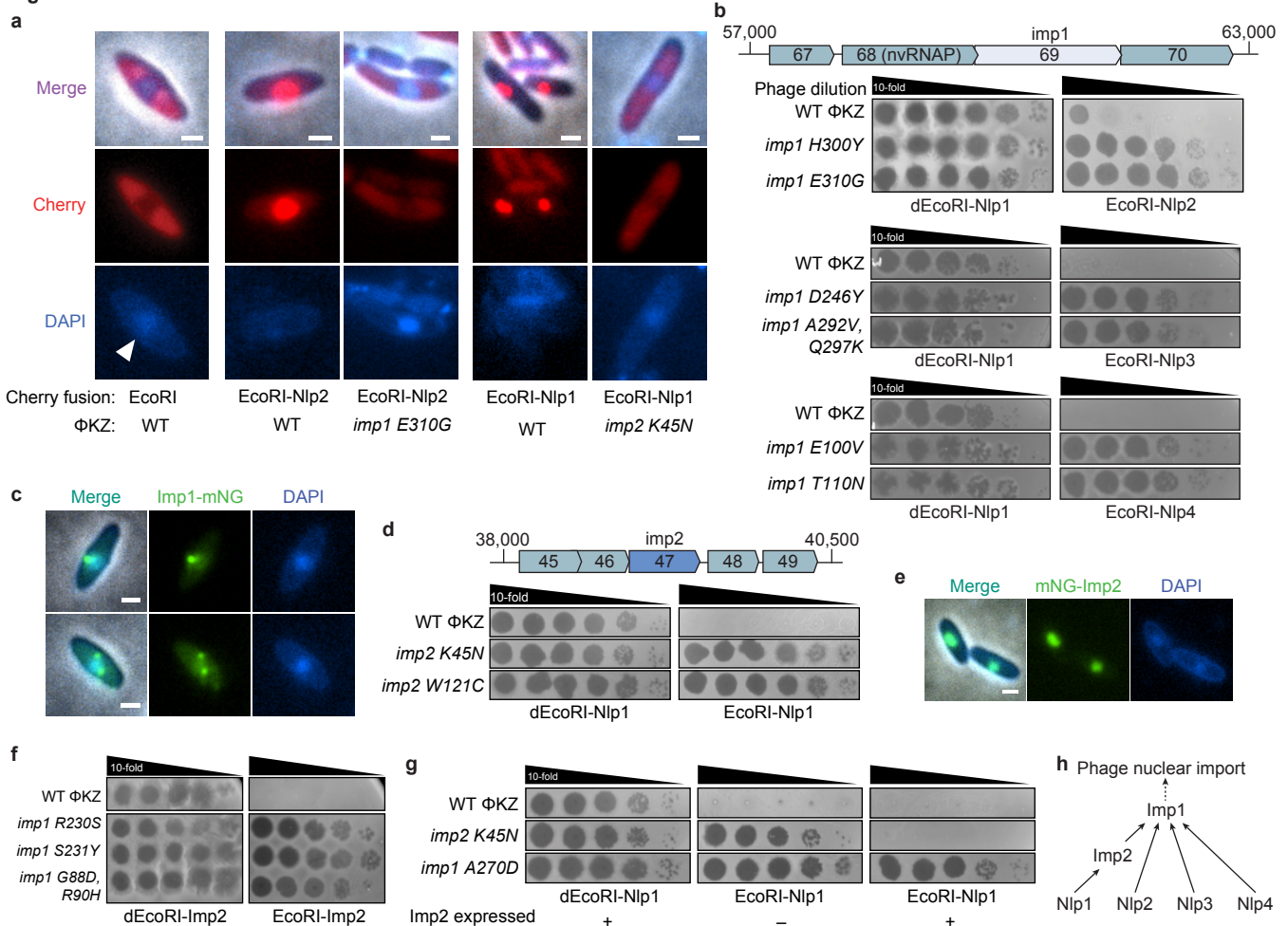
Figure 1: Isolation of mutant phages that reduce protein import into phage nucleus

a, Live-cell fluorescence microscopy of *P. aeruginosa* strain PAO1 expressing the indicated sfCherry2 fusion proteins infected with WT ΦKZ or the indicated phage mutant. DAPI stain indicates phage DNA within the phage nucleus (white arrow). **b**, **d**, **f**, Plaque assays with the indicated WT or mutant phage spotted in 10-fold serial dilutions on a lawn of PAO1 expressing indicated sfCherry2 fusion. dEcoRI, catalytically inactive EcoRI. **c**, **e**, Live-cell fluorescence microscopy of Imp1 fused to mNeonGreen (mNG) infected with WT ΦKZ (n = 94) (**c**) or Imp2 fused to mNeonGreen and infected with WT ΦKZ (n = 129) (**e**). **g**, Plaque assays with the indicated WT or mutant phage spotted in 10-fold serial dilutions on a lawn of PAO1 expressing the indicated sfCherry2 fusion, with or without expression of Imp2 in trans (+/– Imp2). **h**, Schematic representing the factors required for nuclear import of proteins queried in the EcoRI selections, as determined by the escape mutations isolated in *imp1* and *imp2* from each selection (i.e. Nlp2 → Imp1 indicates Nlp2 requires WT Imp1 for its nuclear import). All plaque assays were replicated three or more times. All microscopy experiments were replicated two or more times. Scale bars, 1μm.

Figure 2

a

Phage dilution 10-fold

	dEcoRI-Imp2	EcoRI-Imp2	EcoRI-Nlp2	EcoRI-Nlp3	EcoRI-Nlp4
WT ΦKZ					
R230S					
S231Y					
G88D, R90H					
H300Y					
E310G					
D246Y					
A292V, Q297K					
E100V					
T110N					

c

Phage dilution 10-fold

	dEcoRI-Imp2	EcoRI-Imp2	EcoRI-Nlp2
WT ΦKZ			
R230S			
S231Y			
G88D, R90H			
H300Y			
E310G			
G88D, R90H, **R306L**			
H300Y, **I226T**			

b

Mutations isolated under selection from EcoRI fusion to:
- Imp2
- Nlp2
- Nlp3
- Nlp4

T27A
G88D
E89K
R90H
C121R
M420V
T459N
D246Y
V171F
A292V
Q297K
I226T
T110N
E100G/V
R230C/L/S
S231Y
S232G
F223S
H300Y
M303L
M505K
R306H
E310G

d

- *Klebsiella*
- *Serratia*
- *Pseudomonas*
- *Escherichia/Salmonella*
- *Xanthomonas*
- *Erwinia*
- *Vibrio*
- *Aeromonas*

- No nuclear shell homolog detected with 3xPSI-BLAST or local blastp
- Partial genome record, no nuclear shell homolog detected
- PhuZ homolog detected by blastp, no nuclear shell homolog detected

Tree scale: 1

Figure 2: Imp1 possesses distinct interfaces to enable import specificity

a, **c**, Plaque assays with indicated WT or *imp1* mutant phages spotted in 10-fold serial dilutions on a lawn of PAO1 expressing the indicated sfCherry2 fusion. dEcoRI, catalytically inactive EcoRI. Plaque assays were conducted as in Fig 1b. **b**, AlphaFold2 high confidence predicted structural model of Imp1, with mutated residues from isolated *imp1* mutant phages color coded by the EcoRI selection from which they were isolated. **d**, Phylogenetic tree of Imp1 homologs. A closed black circle indicates that no phage nuclear shell homolog was detected in the corresponding phage genome by 3 iterations of PSI-BLAST or local blastp. Open black circles indicate that the associated genome record is incomplete, and no phage nuclear shell homolog was detected by the same methods. A closed blue circle indicates that no nuclear shell homolog was detected by the same methods, but the tubulin homolog responsible for centering the phage nucleus in the cell, PhuZ, was detected by blastp. All plaque assays were replicated three or more times.
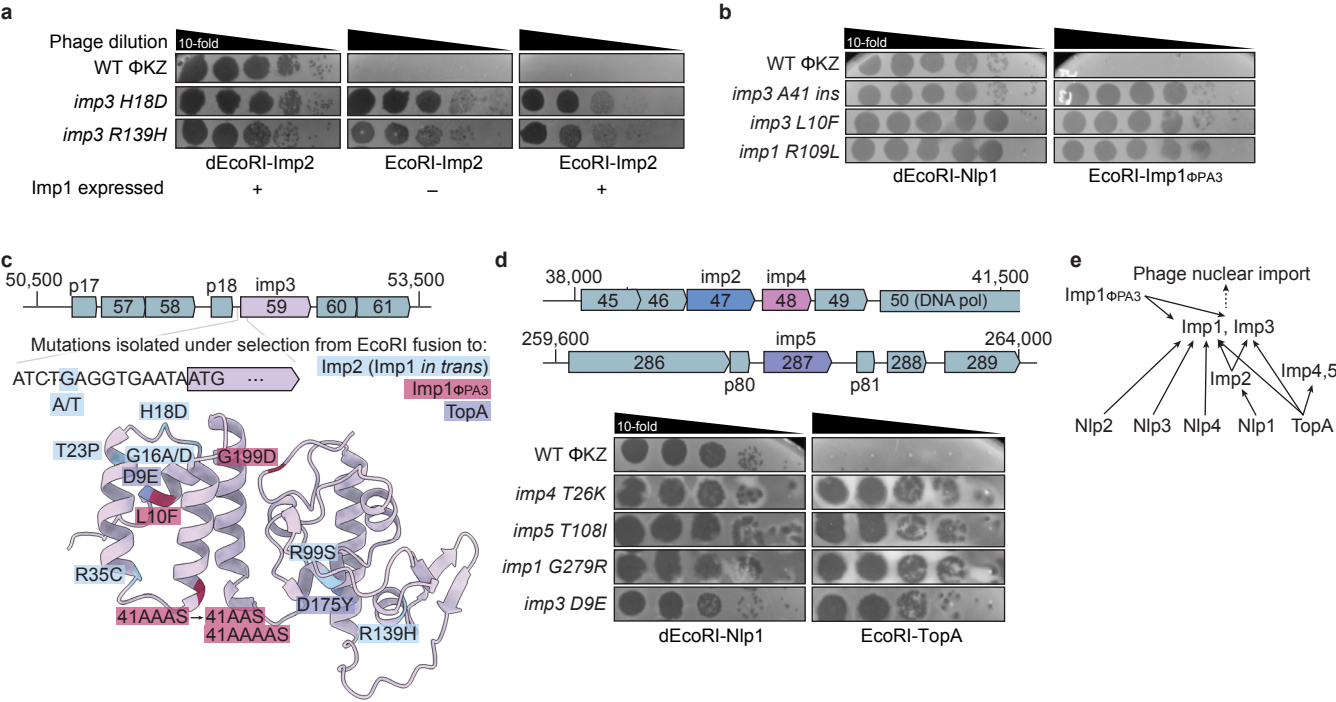
Figure 3

Figure 3: Imp3 is required for proper Imp1 function

a, **b**, Plaque assays with the indicated WT or mutant phage spotted on PAO1 expressing the indicated sfCherry2 fusion, with or without Imp1 expressed in trans (+/– Imp1) (**a**); Imp1ΦPA3 indicates the phage ΦPA3 Imp1 homolog (**b**); dEcoRI, catalytically inactive EcoRI. Plaque assays were performed as in Fig 1b. **c**, AlphaFold2 high confidence predicted structural model of Imp3 with mutated DNA bases or residues from isolated *imp3* mutant phages color coded by the EcoRI selection from which they were isolated. **d**, Plaque assays with the indicated WT or mutant phage spotted on PAO1 expressing the indicated sfCherry2 fusion. **e**, Schematic representing the factors required for nuclear import of all proteins queried in the EcoRI selections in this work, as determined by the escape mutations isolated in *imp1-imp5* from each selection. All plaque assays were replicated three or more times.
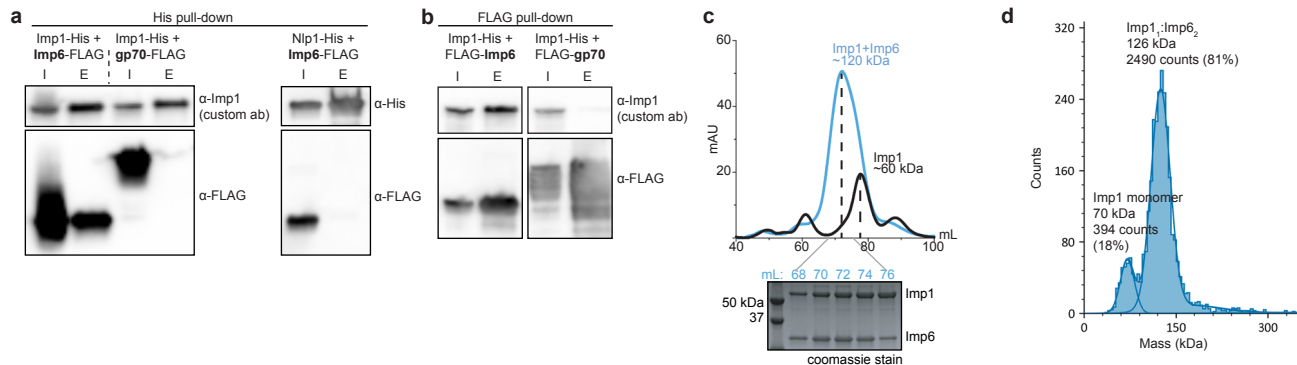
Figure 4

Figure 4: Imp6 stably binds to Imp1

a, **b**, Western blot analysis (anti-Imp1) of a His pull-down of His-tagged Imp1 with FLAG-tagged Imp6, gp70, or Nlp1-His with Imp6-FLAG as a negative control (**a**) or a FLAG pull-down of Imp1-His with FLAG-Imp6 or FLAG-gp70 (right) (**b**); I, input; E, elution. **c**, Size exclusion chromatography (SEC) of Imp1-His alone or in complex with Imp6-FLAG. Peak Imp1-His/Imp6-FLAG SEC fractions (indicated by mL collected off the column) were run on an SDS-PAGE gel and stained with coomassie. **d**, Estimation of Imp1-His/Imp6-FLAG complex molecular weight by mass photometry.