Quantitative Selection Analysis of Bacteriophage φcbK Susceptibility in Caulobacter crescentus

Author(s):
Christen, Matthias; Beusch, Christian; Bösch, Yvonne; Cerletti, Dario; Flores-Tinoco, Carlos E.; Del Medico, Luca; Tschan, Flavia; Christen, Beat

Publication date:
2016-01-29

Permanent link:
https://doi.org/10.3929/ethz-b-000113399

Originally published in:
Journal of Molecular Biology 428(2), https://doi.org/10.1016/j.jmb.2015.11.018

Featured Article

Quantitative Selection Analysis of Bacteriophage φCbK Susceptibility in *Caulobacter crescentus*

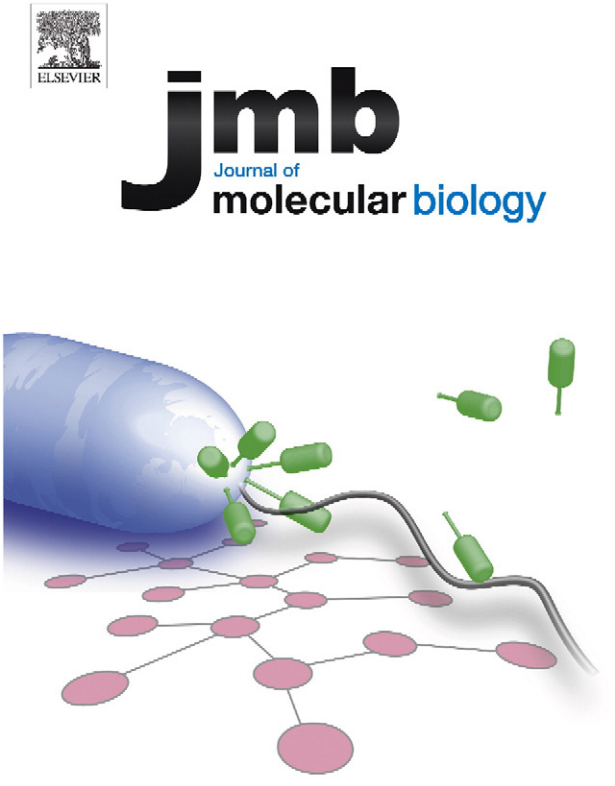
Matthias Christen [1], Christian Beusch [1], Yvonne Bösch [1], Dario Cerletti [1], Carlos Eduardo Flores-Tinoco [1,2], Luca Del Medico [1], Flavia Tschan [1] and Beat Christen [1]

*1 - **Institute of Molecular Systems Biology,** Eidgenössische Technische Hochschule (ETH) Zürich, CH-8093 Zürich, Switzerland*
*2 - **Life Science Zürich PhD Program on Systems Biology,** Zürich, Switzerland*

Correspondence to Beat Christen: *beat.christen@imsb.biol.ethz.ch*
http://dx.doi.org/10.1016/j.jmb.2015.11.018
Edited by M. Gottesman

Abstract

Classical molecular genetics uses stringent selective conditions to identify mutants with distinct phenotypic responses. Mutations giving rise to less pronounced phenotypes are often missed. However, to gain systems-level insights into complex genetic interaction networks requires genome-wide assignment of quantitative phenotypic traits. In this paper, we present a quantitative selection approach coupled with transposon sequencing (QS-TnSeq) to globally identify the cellular components that orchestrate susceptibility of the cell cycle model bacterium *Caulobacter crescentus* toward bacteriophage φCbK infection. We found that 135 genes representing 3.30% of the *Caulobacter* genome exhibit significant accumulation of transposon insertions upon φCbK selection. More than 85% thereof consist of new factors not previously associated with phage φCbK susceptibility. Using hierarchical clustering of dose-dependent TnSeq datasets, we grouped these genes into functional modules that correlate with different stages of the φCbK infection process. We assign φCbK susceptibility to eight new genes that represent novel components of the pilus secretion machinery. Further, we demonstrate that, from 86 motility genes, only seven genes encoding structural and regulatory components of the flagellar hook increase phage resistance when disrupted by transposons, suggesting a link between flagellar hook assembly and pili biogenesis. In addition, we observe high recovery of Tn*5* insertions within regulatory sequences of the genes encoding the essential NADH:ubiquinone oxidoreductase complex indicating that intact proton motive force is crucial for effective phage propagation. In sum, QS-TnSeq is broadly applicable to perform quantitative and genome-wide systems-genetics analysis of complex phenotypic traits.



Legend: *Selective pressure constantly drives bacteriophages to evolve more effective infection strategies such as interlinking their infection and replication cycle to essential host cell components that cannot be disrupted without dramatic fitness costs. The authors used a quantitative selection approach in conjunction with transposon sequencing (QS-TnSeq) to globally identify the cellular networks that orchestrate susceptibility of the cell cycle model bacterium Caulobacter crescentus toward bacteriophage infection. Artwork: Beat Christen and Matthias Christen.*

Introduction

For ages, genomes of bacteriophages and their host cells have undergone adaptive processes. In this arms race, constant selective pressure drives bacteriophages to evolve more effective infection strategies such as interlinking their infection and replication cycle to essential host cell components that cannot be depleted without dramatic fitness costs. The *Caulobacter crescentus* (*Caulobacter* thereafter) bacteriophage φCbK [1] represents an exquisite model system to investigate how phages hijack the complex genetic interaction network of the host cell. Previous genetic studies on φCbK phage susceptibility focused on single genes [2,3] or investigated the effects of distinct protein domains within known regulators of bacteriophage resistance [4]. Although these studies identified pili and flagellar filaments as key components involved in φCbK adsorption, there is still limited information about the genetic network that confers φCbK susceptibility. Genome-wide identification of interaction partners and regulatory pathways that orchestrate φCbK susceptibility is more challenging. With classical molecular genetic approaches, it is difficult to determine the complete number of genes that contribute to a particular phenotype because stringent selective conditions only identify components with dramatic phenotypic responses. However, in the case of complex biological traits such as phage resistance, the phenotypic response is not a binary readout because most bacterial genes involved in phage infection will cause only subtle effects in phage replication [5].

The infection cycle of φCbK contains three consecutive steps (Fig. 1a). Flagella and pili serve as initial points of adsorption [6]. After binding of phage particles to host receptors localized at the flagellated pole of swarmer cells (Fig. 1b), φCbK injects its viral DNA into the host cell in a series of consecutive events that likely involve structural components of type IV pili and its associated secretion machinery [6]. Once the φCbK DNA enters the cytosol, it redirects the cellular machinery toward

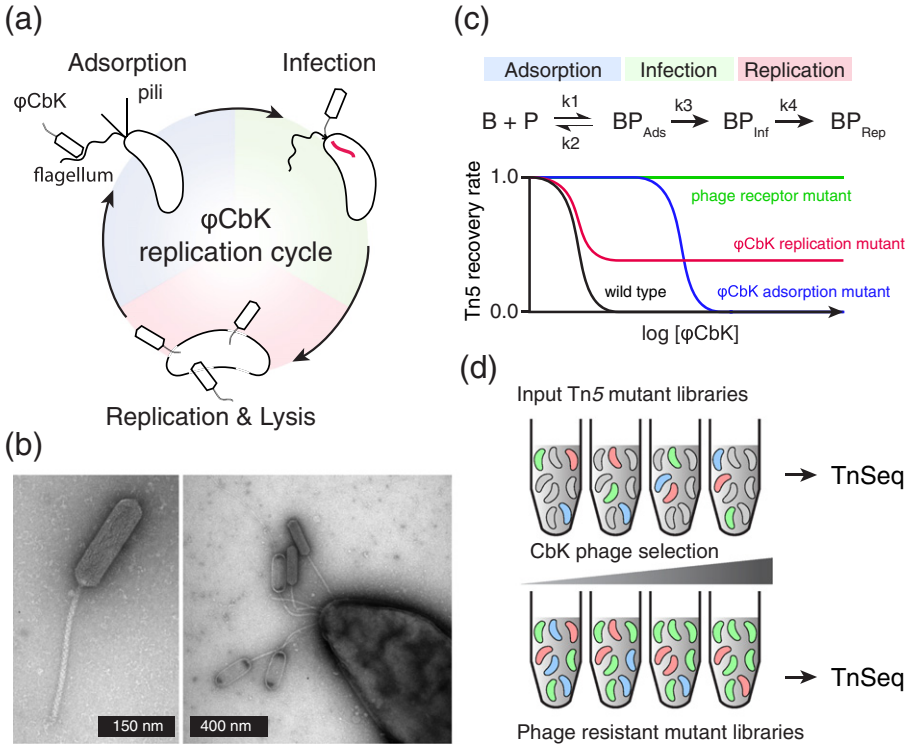


Fig. 1. Replication cycle of the pili-specific bacteriophage φCbK. (a) The double-stranded DNA bacteriophage φCbK adsorbs to the polar flagellum and pili. Upon infection, φCbK injects its genome into the host cell and reprograms the host machinery to replicate the phage genome and produce new infective phage particles that become liberated upon lysis of the host cell. (b) Electron microscopy image of the φCbK phage particles is shown (left panel). φCbK binds to the flagellated cell pole of *Caulobacter* host cells (right panel). (c) Host cell mutants that impair phage adsorption, infection or replication can become resistant against φCbK. The different mutant classes display different host cell survival rates depending on the concentration of phage particles applied. (d) Schema of the quantitative phage selection assay used to challenge a set of six Tn*5* mutant libraries with increasing concentration of bacteriophage titers. Genomic insertion sites prior and after applying phage selection are mapped using transposon sequencing (TnSeq).

phage replication, subsequently leading to the assembly and release of new phage particles. *Caulobacter* mutants with impairments in phage adsorption, infection or replication become resistant to φCbK to varying degrees. Based on a simplified kinetic model of phage infection, we reasoned that, depending on the concentration of phage particles applied during a selection experiment, different mutant classes display differences in host cell survival rates (Fig. 1c and d). In this model, φCbK phage particles (P) associate reversibly with bacteria (B). Subsequently, they inject their DNA in a nonreversible process into the host cell and provoke phage replication. Mutations in genes involved in initial adsorption will exhibit a decreased binding affinity for phage particles. This mutant class regains phage sensitivity under increased concentrations of phage (Fig. 1c, blue line). Mutations in structural components of the phage receptors and their associated transport channels, which mediate phage infection, will show a resistance profile independent of the phage concentration applied. Such receptor mutants remain phage resistant even in the presence of high phage titer (Fig. 1c, green line). Mutants impaired in phage replication have identical adsorption and infection rates as wild-type cells. However, because efficiency of phage replication is diminished, this mutant class exhibits increased survival rates compared to those of wild-type cells (Fig. 1c, red line). Based on this model, we speculated that dose-dependent selection experiments measuring recovery rates of transposon insertions as a function of phage concentration will reveal systems-wide gene sets of *Caulobacter* that are required for the adsorption, infection and replication stages of bacteriophage CbK.

Results and Discussion

Identification of φCbK susceptibility genes using quantitative selection analysis coupled to transposon sequencing

To identify the complete set of genes required for successful adsorption, infection and φCbK replication, we designed a systems-wide forward genetic screen based on quantitative selection analysis coupled to transposon sequencing (QS-TnSeq) (Fig. 1d). Hypersaturated Tn*5* transposon mutagenesis is used to generate a large pool of single insertion mutants. During subsequent phage selection, mutants harboring Tn*5* insertions within genes not related to φCbK infection are lost from the mutant pool. Cells harboring insertions within genes involved in phage adsorption, infection or replication exhibit higher probabilities to persist in the presence of phages and thus will remain in the mutant pool. A total of six independent transposon insertion mutant libraries were generated using our previously described mutagenesis procedure [7]. For effective sample multiplexing, we used a modified Tn*5* transposon with internal barcode sequences to tag individual mutant pools (see Materials and Methods and Table S3). Next, we challenged the six different input mutant pools separately with increasing titers of phage φCbK using a multiplicity of infection (MOI) of 10^{-1} to 10^{4} bacteriophages per bacterial cell (Table 1, Table S1 and Materials and Methods). To detect changes in transposon mutant composition upon phage selection, we mapped genome-wide insertion sites from the input and output mutant pools using a previously established transposon sequencing (TnSeq) strategy [7]. From the six input mutant pools, we recovered between 3.0×10^{6} and 3.6×10^{6} paired-end reads that unambiguously mapped to the *Caulobacter* NA1000 genome and subsequently assigned between 64,810 and 82,674 unique Tn*5* insertions per library (Table 1 and Table S1). Using a custom sequence analysis pipeline, we analyzed global insertion statistics and calculated transposon insertion occurrence and distributions within each annotated GenBank feature of the *Caulobacter* NA1000 genome (Supplementary Data, Table S5 and Materials and Methods).

Global mapping of transposon insertion sites

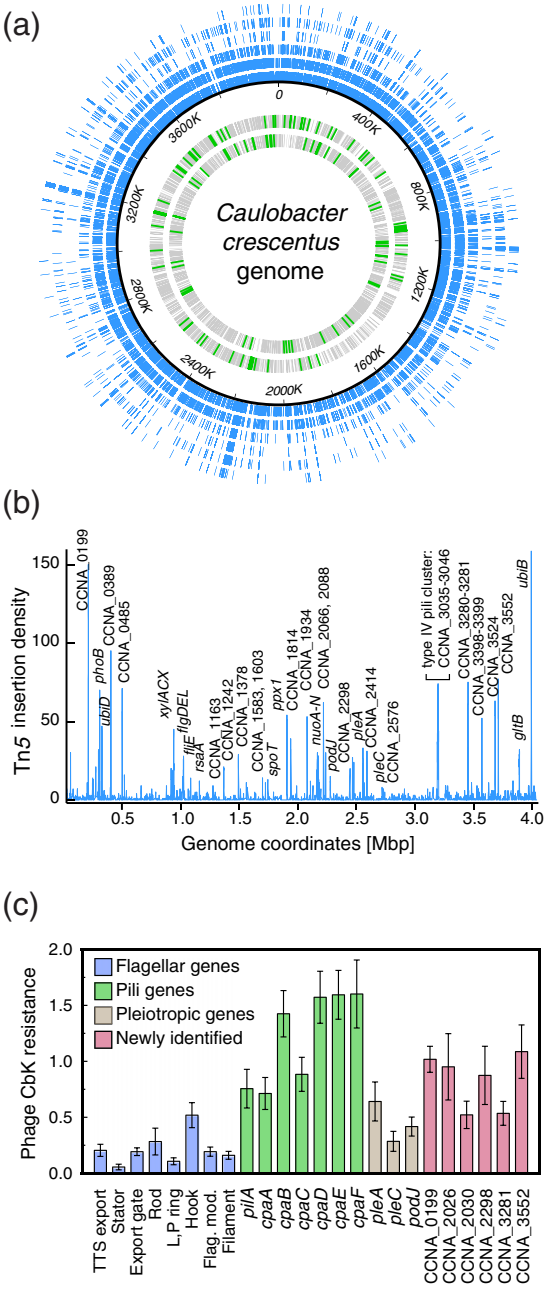
We assessed the Tn*5* recovery rate across the 4-Mb *Caulobacter* genome upon selection with the φCbK. As a general tendency, we observed that the genome-wide frequency of Tn*5* insertion decreased in the presence of increasing phage concentration (Fig. 2a, Table 1 and Table S2). Under the least selective condition (MOI of 10^{-1}), we recovered 18,336 unique Tn*5* insertions while only 1455 Tn*5* insertions were recovered for the most stringent selection conditions (MOI of 10^{4}) (Table 1 and Table S2). For each of the 3885 annotated protein-coding genes of *Caulobacter*, we determined the number of

Table 1. Phage concentration-dependent recovery of phage-resistant Tn*5* mutants

Mutant library	Selection (MOI)[a]	Unique Tn*5* hits[b]	Phage susceptibility genes[c]
Tn*5* pool 1	10^{4}	1455	58
Tn*5* pool 2	10^{3}	4064	74
Tn*5* pool 3	10^{2}	2279	66
Tn*5* pool 4	10^{1}	6183	87
Tn*5* pool 5	10^{0}	13,376	103
Tn*5* pool 6	10^{-1}	18,336	130
Total		41,108	208

[a] MOI specifies the ratio of bacteriophages to bacteria applied for φCbK selection of the different Tn*5* mutant libraries.
[b] Estimated number of unique transposon insertions represented within the different Tn*5* mutant pools.
[c] Host genes contributing to the susceptibility against φCbK.

transposon insertion recovered from the input and output mutant pools across the six phage selection conditions applied (Supplementary Data, Table S5 and Materials and Methods). Assuming random recovery of transposon mutants, we calculated for every gene the probability (*P*-value) of recovering equal or higher numbers of transposon insertions after phage selection (see Materials and Methods, Supplementary Data and Tables S6 and S7). Genomic regions where the insertion frequency remains high under elevated selective pressure pinpoint to key genes important for phage susceptibility. A set of 135 genes representing 3.30% of the *Caulobacter* genome exhibited significant accumulation of transposon insertions ($P < 10^{-5}$) upon φCbK selection (Table S5 and Fig. S1). We observed that phage susceptibility genes are distributed throughout the *Caulobacter* genome with some gene sets grouping together at distinct genomic loci (Fig. 2a and b). Functional classification among COG categories revealed that 14 genes (10.4%) are involved in energy production and conversion, 11 genes (8.2%) function in amino acid metabolism and transport and 10 genes (7.4%) participate in transcription. An additional 9 genes (6.7%) are involved in cell motility and 8 genes (5.9%) are related to intracellular transport and secretion. Surprisingly, 28 genes representing 20.7% of the identified φCbK susceptibility genes are currently of unknown function and have not been previously implicated in φCbK infection. Furthermore, we observed significant accumulation of φCbK susceptibility genes within the COG categories of "intracellular transport and secretion" (*P*-value of 4.2×10^{-4}), "energy production and conversion" (*P*-value of 8.9×10^{-4}) and "cell motility" (*P*-value of 1.2×10^{-2}), which suggest that these three COG subgroups contain modules with important roles for phage propagation.



Discovery of novel type IV pili genes required for φCbK propagation

Among the hypersusceptibility genomic regions, we discovered the highest density of Tn*5* insertions

Fig. 2. Genome-wide mapping of transposon mutants conferring φCbK resistance. (a) A total of 41108 unique Tn*5* insertions sites (blue) were recovered after challenging an input Tn*5* mutant library, consisting of 394935 unique transposon insertion mutants, with increasing concentrations of φCbK. Transposon recovery rate decreases with increasing bacteriophage concentration (blue outer circles, strength of phage selection is increasing for tracks further outwards). The 4-Mb *Caulobacter* genome encompasses 3885 annotated protein-coding sequences (CDS) shown in the inner (minus strand) and outer (plus strand) gray tracks. Cumulatively, 135 CDS (green marks) associated with phage susceptibility were identified and display a high transposon recovery rate upon phage selection. (b)The cumulative transposon insertion density for all six phage selections is plotted across the 4-Mb *Caulobacter* genome using a 1-kb sliding window. Genes and gene clusters showing high transposon insertion densities are labeled by the corresponding allele names or the *Caulobacter* locus tag. (c) The quantified phage φCbK resistance levels upon disruption of genes involved in flagellar assembly (blue), pili biogenesis (green), pleiotropic organelle development (brown) and the newly identified phage susceptibility genes (red) are shown. Phage φCbK resistance is calculated as the ratio between the number of transposon insertions upon phage selection and the initial number of transposon insertions present within the input mutant pool.

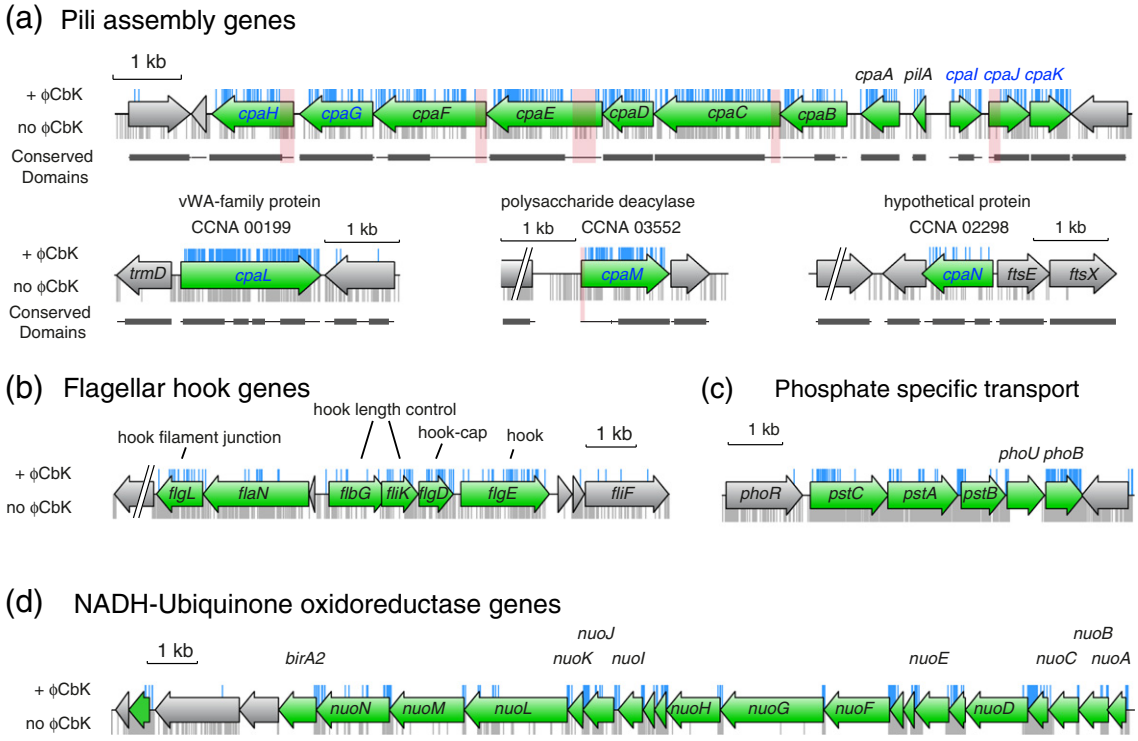


Fig. 3. Identification of genes and gene clusters required for φCbK propagation. (a) Within the *Caulobacter* type IV pili gene cluster, a high number of transposon insertions (blue marks) are recovered after phage selection of the input transposon mutant pool (gray marks). Genes (arrows) that are required for bacteriophage propagation are highlighted in green. The newly identified pili genes *cpaH*, *cpaG* and *cpaI*, *cpaJ* and *cpaH* (blue label) are adjacent to the previously identified *Caulobacter* pili genes *pilA*, *cpaA*, *cpaB*, *cpaC*, *cpaD*, *cpaE* and *cpaF*. Three additional genes *cpaL* (CCNA_00199), *cpaM* (CCNA_03552) and *cpaN* (CCNA_02298) residing outside of the main pili gene cluster show high transposon recovery rate upon phage selection, indicating that they are likely to be involved in pili biogenesis. Black bars below the genome annotation track indicate conserved protein domains as defined by the National Center for Biotechnology Information conserved domains database [25]. Absence of transposon insertions within the 5′ region of the pili genes *cpaE*, *cpaF*, *cpaH* and *cpaJ* (highlighted in red) identifies CDS that are shorter than listed in the current *Caulobacter* NA1000 genome annotation. (b) Transposon insertion data are plotted prior to (gray) and after phage selection (blue) for flagellar hook genes *flgL*, *flaN*, *flbG*, *fliK*, *flgD* and *flgE* (green) that display a high transposon recovery rate compared to other structural flagella genes such as *fliF*. (c) Transposon insertion data are shown prior to (gray) and after phage selection (blue) for the phosphate-specific transport genes. Transposon insertions within *pstC*, *pstA*, *pstB* and *phoB* (green) confer φCbK phage resistance. (d) Transposon insertion data are plotted prior to (gray) and after phage selection (blue) for the proton-pumping NADH:ubiquinone oxidoreductase respiratory complex I of *Caulobacter*. A set of 14 *nuo* genes form a complex that couples electron transfer from NADH to ubiquinone with the translocation of protons across the membrane.

within the *Caulobacter* pili gene cluster *cpaA-F* encoding several components of the type IV pili secretion apparatus (Figs. 2c and 3a). In addition to the five genes *cpaA-F*, which have previously been implicated in *Caulobacter* pili biogenesis and φCbK phage resistance, we recovered Tn*5* insertions within the two coding sequences (CDS) located downstream of the pilus assembly ATPase *cpaF*. CCNA_03036 and CCNA_03035 share homologies with *tadB*- and *tadC*-related pilus assembly genes from *Aggregatibacter actinomycetemcomitans*. In addition, we identified three CDS upstream of the pilin gene *pilA* that exhibit high density of Tn*5* insertions. CCNA_03044 is annotated as a CpaC-related secretion pathway protein, and CCNA_03045 and CCNA_03046 encode TadG- and TadE-related pilus assembly proteins. Furthermore, we found similar transposon insertion recovery rates for three orphan genes (CCNA_00199, CCNA_02298 and CCNA_03552) located outside of the *cpaA-F* cluster (Fig. 3a). CCNA_00199 shares N-terminal homologies with *tadG*, a flap pilus assembly protein from *A. actinomycetemcomitans* [8]. Named after their involvement in tight adherence phenotype, *tad* genes are predicted to encode structural components for the assembly and secretion of Flp pili. CCNA_00199 shares C-terminal homologies with the Willebrand factor type A domain, which is widely distributed among Bacteria, Archaea and

Eukarya and has been implicated in mediating contacts between the extracellular matrix and cells [9]. CCNA_02298 encodes a protein of unknown function, which contains an N-terminal CXXCX(19)CXXC motif, suggestive of zinc fingers and thioredoxin. CCNA_02298 shares partial homologies to *agmX*, whose interruption affects adventurous gliding motility in *Myxococcus xanthus* [10]. CCNA_03552 contains a C-terminal putative catalytic domain of family 2 polysaccharide deacetylases that presumably removes N-linked acetyl groups from cell wall polysaccharides. We annotated these 8 newly identified phage susceptibility genes as *cpaG* (CCNA_03036), *cpaH* (CCNA_03035), *cpaI* (CCNA_03044), *cpaJ* (CCNA_03045), *cpaK* (CCNA_03046), *cpaL* (CCNA_00199), *cpaM* (CCNA_03552) and *cpaN* (CCNA_02298) (Fig. 3a).

Transposon insertions in flagellar hook genes confer φCbK resistance

During flagellar biogenesis, several unique mechanisms coordinate gene expression with distinct stages of the organelle assembly process [11]. During the *Caulobacter* cell cycle, a single flagellum is built up at the new pole of the predivisional cell. The flagellum filament remains paralyzed until daughter cells become compartmentalized upon cell division. Shortly after cell division and release of the swarmer cell, pili biosynthesis is initiated and the flagellum becomes functional [12]. Several genes of the flagellar biogenesis apparatus have previously been reported to affect efficiency of φCbK attachment [13,14]. We examined Tn*5* insertion frequency within the genome-wide set of motility genes and flagellum biosynthetic genes. Surprisingly, elevated Tn*5* insertion frequencies were exclusively observed within the 7 genes encoding hook-basal body complex (*fliE*, *flgC*), flagellar hook (*flgE*), hook cap (*flgD*), hook length control (*fliK* and *flbG*) and hook filament junction (*flgL*) (Figs. 2c and 3b). However, no significant numbers of Tn*5* insertions were recovered for the 79 remaining COG genes involved in flagellar motility (Fig. 2c, Supplementary Data and Tables S7 and S8). Disruption of most of flagellar genes causes a nonmotile phenotype but apparently does not confer φCbK resistance. Therefore, the lack of a functional flagellum does not explain the φCbK resistance phenotype observed for flagellar hook genes. In the *Caulobacter* predivisional cells, pili assembly is temporally initiated after completion of flagellar biogenesis during a short time window of the cell cycle before activation of flagella filament rotation occurs [15]. It is not well known how biogenesis of the type IV pili export apparatus is temporally linked to the flagellum, but it likely involves checkpoints of the flagellum biogenesis machinery for its initiation. Therefore, completion of flagellar hook assembly may constitute a signal that activates pili biogenesis in *Caulobacter*.

Phosphate-specific transport and NADH:ubiquinone oxidoreductase affect φCbK susceptibility

In addition to pili and flagellar hook genes, we detected high densities of Tn*5* insertions within some of the genes encoding the ABC-type phosphate transport system *pstABCS* and its cognate two-component system *phoR-phoB* (Fig. 3c). In the presence of excess phosphate, the phosphate-specific transmembrane channels PstA, PstC and the ATPase PstB form a repression complex for PhoR [16]. During phosphate starvation, PhoR repression by PstABC is released. Subsequently, PhoR undergoes autophosphorylation and phosphorylates its cognate response regulator PhoB. Within the *pstABC* genes, we observed high transposon recovery rate for the *phoB* gene and the region encoding the N-terminal cytoplasmic regulatory domain of the transmembrane channel PstC. Tn*5* insertions within *phoB* likely impair PhoB-mediated transcriptional control. Because Tn*5* insertions were not recovered within the sensor kinase *phoR* (Fig. 3c), it is unlikely that *phoB*-mediated φCbK resistance is caused by the absence of PhoR-specific phosphotransfer and may involve signaling by nonphosphorylated PhoB or phosphotransfer by a PhoR-independent mechanism. Deletion of *phoB* might affect pili biogenesis by a mechanism similar to its regulatory function in stalk biogenesis [17].

Additional enrichment of Tn*5* insertions upon phage selection was detected within genes of the NADH:ubiquinone oxidoreductase respiratory complex I (Fig. 3d). In *Caulobacter*, a set of 14 *nuo* genes couple electron transfer from NADH to ubiquinone with the translocation of protons across the membrane. Because proton motive force (PMF) generation drives cellular ATP synthesis, many of the *nuo* genes are essential and cannot be disrupted by Tn*5* insertions. However, we observed high Tn*5* density within several promoters and translational regulatory features (Fig. 3d). These mutations likely affect stoichiometry and proper assembly of the multienzyme complex resulting in impaired PMF and reduced ATP synthesis. Studies in *Vibrio cholerae* have shown that mutations within *nuo* genes affect flagellar motility [18]. φCbK uses rotating flagellar filaments for initial contact and movement of phage particles to the cell surface receptors [6]. However, because flagellar genes outside of the hook gene showed no significant phage resistance levels (Fig. 2c and Table S6), it is unlikely that phage tolerance of *nuo* genes arises due to an impairment of flagellar rotation. Therefore, it is more plausible that phage replication and assembly of φCbK phage particles is affected by depleted PMF and cellular ATP levels in a mechanism similar to the one reported for the filamentous phage f1 [19] where ATP hydrolysis and presence of a fully functional

PMF is a prerequisite for successful phage replication and virion assembly.

Protein domains required for φCbK infection

To define essential and nonessential regions of the coding sequences required for φCbK replication, we analyzed the distribution of transposon insertion sites within the 135 identified phage susceptibility genes. We discovered 41 ORFs (open reading frames) that had transposon insertions within a significant portion of their 3′ region but lacked insertions within the 5′ region (Table S9). Conversely, we found 66 ORFs that tolerated disruptive transposon insertions within the 5′ region while no insertion events were tolerated further downstream (Table S10). In-depth analysis of Tn*5* insertion patterns allowed us to detect protein-coding sequences that are likely shorter than annotated in the current *Caulobacter* genome annotation (Fig. 3a) and allowed us to resolve regions within proteins that are critical for φCbK susceptibility (Table S9). For example, transposon insertions conferring phage resistance are nonuniformly distributed along the ORF of *cpaE* (Fig. 4a). The current annotation of the *cpaE* gene includes a nonconserved N-terminal proline-rich domain (P-rich) of unknown function, followed by an atypical receiver domain (REC), an ATPase domain and a C-terminal membrane targeting sequence (MTS) [20,21]. Interestingly, transposon insertions within the N-terminal P-rich domain are lost upon phage selection (Figs. 3a and 4a), suggesting that the presence of the P-rich domain is dispensable for phage infection (Fig. 4a). Transposon insertions within the 3′ end of a gene will abrogate transcription and translation unless the gene is misannotated and includes a transcriptional and translational start site downstream of the insertion site. Absence of recovered transposon insertions within the 5′ portion of *cpaE* suggests that *cpaE* contains a misannotated or a second start site located downstream of the P-rich domain. To test this hypothesis, we analyzed φCbK resistance phenotypes of various N-terminal deletion variants of *cpaE* in a Δ*cpaE* background (Materials and

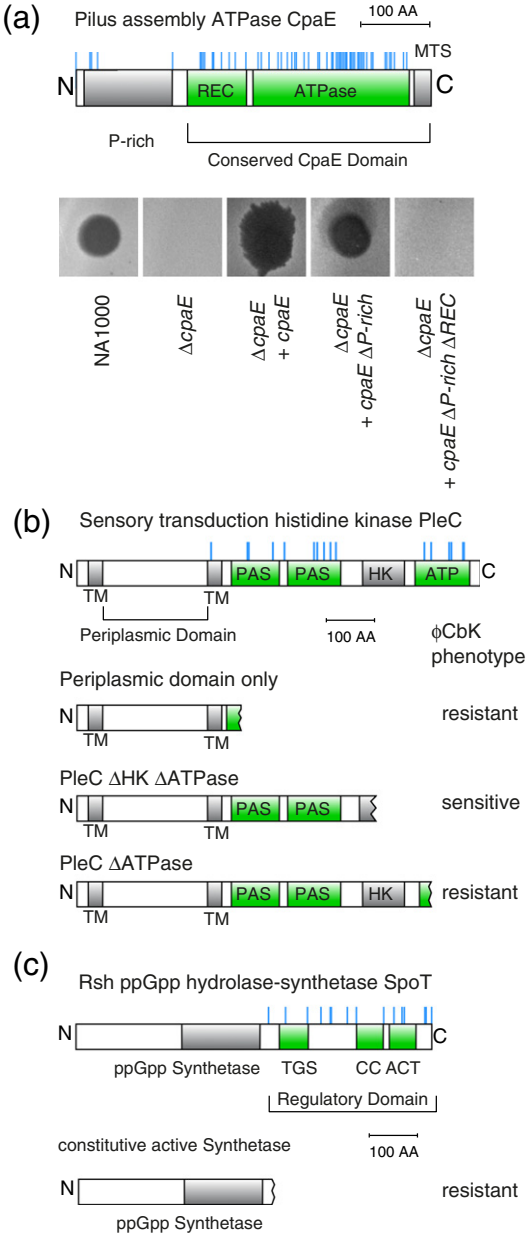


Fig. 4. Identification of protein domains required for phage propagation. (a) Transposon insertions conferring phage resistance (blue) are nonuniformly distributed and predominantly target the conserved atypical receiver and ATPase domain of *cpaE*. Absence of transposon insertions within a 305-bp-long segment encoding a not conserved N-terminal region (100 amino acids long) of low sequence complexity indicates a misannotated start site. Phage-resistant phenotypes from wild type and N-terminal *cpaE* truncations are shown below the top panel as determined by a phage spot assay. (b) The domain architecture and transposon insertion pattern of the sensory transduction histidine kinase PleC is shown. No transposon insertions were recovered within an N-terminal segment (286 amino acids long) spanning the periplasmic domain of *pleC*, as well as within the histidine kinase domain (HK) of PleC. Insertions that disrupt the PAS domain or ATPase domain of PleC are phage resistant. Likewise, mutations that target the C-terminal ATPase domain but leave the HK domain intact are phage resistant. The corresponding truncated protein products and their phage resistance phenotype are shown below the top panel. (c) The SpoT protein contains a C-terminal regulatory domain that inhibits the activity of the N-terminal ppGpp synthase domain. Transposon insertions within the C-terminal regulatory domain of SpoT result in a constitutive active the N-terminal ppGpp synthase [23] indicating that elevated levels of ppGpp elicit phage resistance.

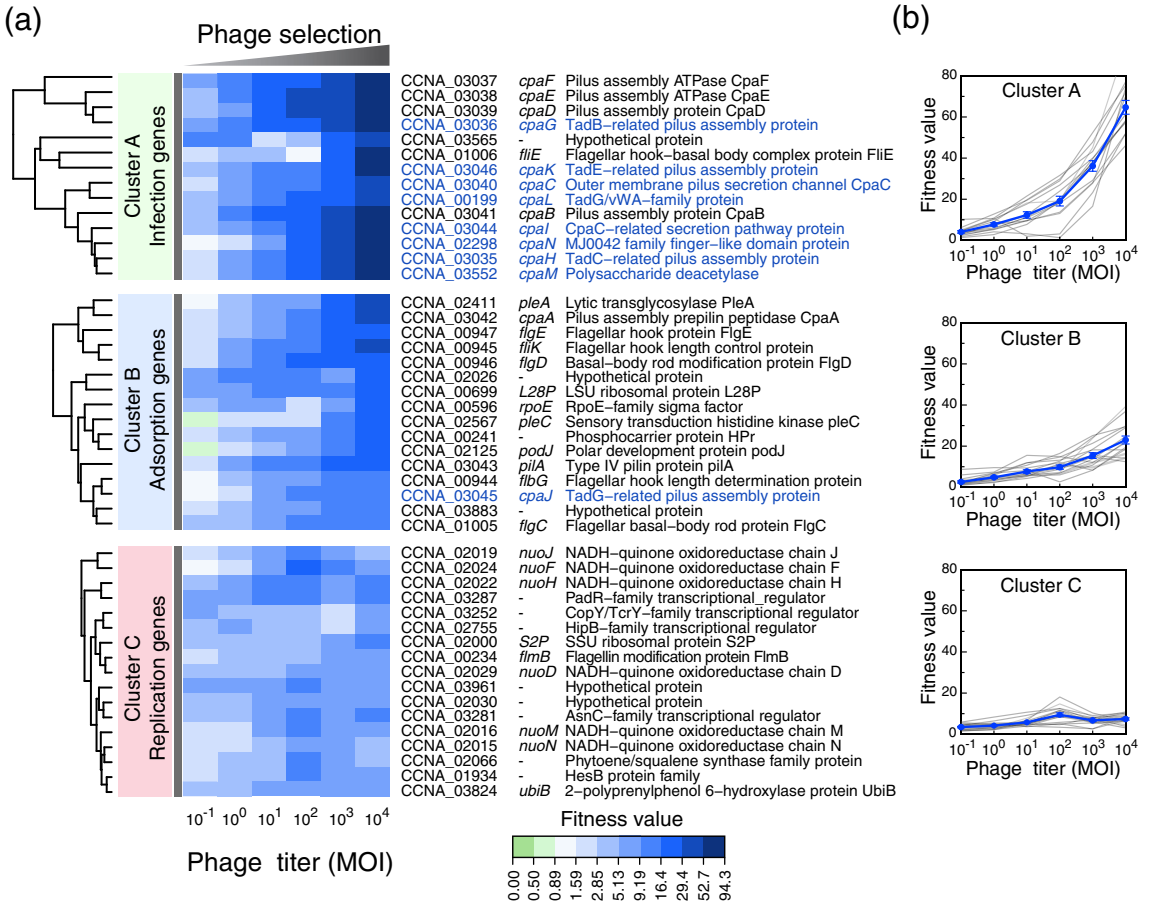


Fig. 5. Hierarchical clustering analysis identifies genetic modules involved in bacteriophage CbK susceptibility. (a) A set of 135 genes that showed significant enrichment of transposon insertion upon phage selection ($P < 10^{-5}$) were selected for hierarchical clustering analysis. For every gene, the calculated logarithmic fitness value for increasing φCbK phage titers (MOI) is represented as a heat map. Clustering was performed along the genotype axis using Euclidean pairwise distances and linkage based on the average distance between clusters. Dendrograms representing the linkage of mutant alleles within the identified clusters A, B and C are shown. Genes in cluster A exhibit the largest increase in fitness upon φCbK infection and contain known and newly identified components (in blue type) of the type 4 pili secretion machinery. Pleiotropic genes, components involved in peptidoglycan remodeling, polar protein localization factors as well as flagella hook and pili subunit genes are found in cluster B. Cluster C contains genes encoding components from proton-pumping NADH:ubiquinone oxidoreductase respiratory complex I and factors involved in flagellin modification. The complete hierarchical clustering of all 135 genes analyzed is provided in Fig. S2. (b) Average fitness (in blue) and fitness graphs for individual genes (in gray) are plotted for clusters A, B and C.

Methods). Deletion of the chromosomal *cpaE* genes results in phage-resistant phenotype while plasmid-based expression of CpaE in a Δ*cpaE* background complements the phage-resistant phenotype (Fig. 4a). We found that a truncated version of *cpaE* lacking the P-rich domain reverted the phage-resistant phenotype of a *cpaE* deletion mutant while an N-terminal *cpaE* deletion spanning the P-rich domain and adjacent REC domain did not revert the resistance phenotype (Fig. 4a). These findings demonstrate that the presence of a functional REC, ATPase and MTS domain is sufficient for CpaE to mediate phage susceptibility.

Similar nonuniform Tn*5* insertion patterns were observed for the histidine kinase gene *pleC*, where transposon insertions were absent within the histidine kinase domain and an N-terminal segment (286 amino acids long) spanning the periplasmic domain of PleC (Fig. 4b). Therefore, truncated versions of the periplasmic domain remain φCbK sensitive while presence of an intact periplasmic domain without adjacent PAS domains causes phage resistance. These observations suggest that the periplasmic part of *pleC* interacts directly or indirectly in a PAS-dependent manner with a factor required for phage susceptibility. Similarly, our finding that insertions within the C-terminal ATPase domain are prevalent but insertions within the histidine kinase domain are absent suggests that a structurally intact but catalytically inactive histidine kinase domain confers CbK phage resistant.

A nonuniform distribution of transposon insertions was also observed within the coding sequence of the *spoT* gene (Fig. 4c and Fig. S1). SpoT synthesizes the stringent response messenger (p)ppGpp in the presence of stalled ribosomes, carbon or nitrogen shortage [22]. The C-terminal TGS/CC/ACT domains of SpoT inhibit ppGpp synthesis of the N-terminal catalytic domain under optimal growth conditions [23]. Tn*5* insertions within the regulatory region of *spoT* likely result in constitutively activated SpoT. Elevated levels of ppGpp have been shown to delay swarmer to stalk cell transition and chromosome replication [23]. In addition to these known phenotypic effects, our results suggest that ppGpp is involved in modulating the efficiency of φCbK infection.

Hierarchical clustering analysis of selection-dependent gene fitness datasets

To define the genetic modules that confer adsorption, infection and phage replication, we analyzed Tn*5* recovery rates of transposon insertions as a function of phage concentration. For each *Caulobacter* gene and selection condition, we calculated a fitness value based on the recovered transposon insertion density (Supplementary Data, Table S6 and Materials and Methods). Next, we performed hierarchical clustering of selection-dependent fitness values of a total of 135 *Caulobacter* genes that showed significant enrichment ($P < 10^{-5}$) of transposon insertion upon phage selection (Supplementary Data, Table S6 and Fig. S2). We identified three gene clusters (clusters A, B and C) that display a high transposon recovery rate upon phage selection (Fig. 5a and b), as well as three additional clusters (clusters D, E and F) with genes displaying less pronounced insertion sites enrichment (Fig. S2). Cluster A encompasses 14 genes that share a strong upshift in fitness in the presence of high CbK phage titers. Because persistence of phage resistance is expected for mutants that lack functional phage receptors, it is likely that these 14 genes constitute the core receptor complex that enables binding of φCbK to the cell surface and translocation of its DNA into the host cell cytoplasm. In agreement with this hypothesis, cluster A contains well-characterized genes of the pili biogenesis operon including the ATPases *cpaE*, *cpaF*, the outer membrane pilus secretion channel *cpaC* and *cpaB*, *cpaD*. In addition, the flagellar hook-basal body complex *fliE* and eight genes, not previously reported to confer φCbK resistance, were classified within cluster A. Interestingly, several of these genes such as the newly annotated genes *cpaG*, CCNA_03565, *cpaK*, *cpaL*, *cpaN*, *cpaH* and *cpaM* phenotypically clustered together with the characterized pili biogenesis genes, suggesting a common role in φCbK resistance mechanisms and their potential involvement in pili biogenesis.

Cluster B contains the pleiotropic genes *pleA*, *pleC* and *podJ*, as well as structural components of the flagellar hook, the pilus filament *pilA*, an uncharacterized RpoE-like sigma 70 factor (CCNA_00596) and two hypothetical proteins (Fig. 5a). Genes encompassing cluster B show a lower degree of phage resistance in the presence of high phage titers compared to genes present in cluster A. This suggests that φCbK infection through surface receptors still partially occurs when genes from cluster B are disrupted by transposons. Interestingly, the lytic transglycosylase *pleA* phenotypically groups with the prepilin peptidase *cpaA*, the flagellar rod modification gene *flgD*, the hook gene *flgE* and the hook length control gene (*fliK*), suggesting a potential genetic submodule that affects pili biogenesis upon regulation of *cpaA*. The two pleiotropic genes *pleC* and *podJ* form a second subcluster with *pilA*, *cpaJ*, *flbC* and *flbG*. Interestingly, deletions of most genes from cluster B cause defects in both flagellar and pili biogenesis. A *pleC* deletion for example harbors a paralyzed flagellum and *podJ*-null mutants are unable to form pili [24]. We concluded that cluster B predominantly encompasses genes affecting biogenesis of extracellular organelle structures involved in initial phage adsorption.

Cluster C encompasses 17 genes whose fitness values are moderate but nevertheless remain constant over the different phage concentrations applied (Fig. 5a and b). According to our phage infection model (Fig. 1c), such fitness profiles are expected for host genes involved in the latent period of the φCbK replication cycle. In agreement with our model, we found that cluster C includes six genes encoding for structural components of the NADH:ubiquinone oxidoreductase respiratory complex, a HesB family protein involved in synthesis of iron-sulfur cofactors, a putative squalen synthase family protein (CCNA_02066) and *ubiB*, the 2-polyprenylphenol 6-hydroxylase encoding gene of the ubiquinone biosynthetic pathway. These genes are likely critical for keeping the host cell energized, which is a prerequisite for effective phage particle reproduction. Similarly, the essential ribosomal protein S2P (CCNA_02000) was also included in cluster C. In addition, we found a set of four putative transcriptional regulators from the PadR (CCNA_03287), CopY (CCNA_03252), HibB (CCNA_02755) and AsnC (CCNA_03281) family (Fig. 5a and Table S4); two hypothetical proteins of unknown function (CCNA_02030 and CCNA_03961); and the flagellin modification gene *flmB* within cluster C. Whether the four genes encoding transcription factors are involved in maintenance of the metabolic state of the host cell or are directly engaged in modulating the transcription of phage genes remains to be determined.

Phenotypic characterization of the pili genes *cpaL* and *cpaM*

Results from the hierarchical clustering suggested that the two genes *cpaL* and *cpaM* are principal determinants of φCbK susceptibility and thus likely represent components of the pili biogenesis machinery. We performed transposon mutagenesis and isolated a set of 288 individual φCbK-resistant clones (Materials and Methods). Among the mutants obtained, we selected two Tn*5* insertions within *cpaL* and two *cpaM* for further phenotypic characterization. Upon hierarchical clustering, the *cpaL* and *cpaM* genes group phenotypically in close proximity to known pili assembly genes, suggesting potential involvement of *cpaL* and *cpaM* in *Caulobacter* pili biogenesis. However, φCbK resistance might not be caused exclusively by defects in pili biogenesis since mutations with impaired flagella also exhibit some impairment in φCbK adsorption and hence a reduction in φCbK susceptibility [6]. To investigate whether impaired flagellar biogenesis or function contributes to φCbK resistance of *cpaL* and *cpaM*, we assessed the motility phenotypes (Materials and Methods). A motility assay on 0.35% swarming agar revealed that neither *cpaL* nor *cpaM* disruptions abolished motility. While Tn*5* insertions within *cpaL* resulted in swarming sizes similar to those of the wild type, insertions within *cpaM* resulted in much smaller swarms (Fig. S3a). Bacteriophage φCbK has been shown to preferentially adsorb to the swarmer cell type and to use the flagellar filament and polar pili for a two-step contact mechanism [6]. Electron microscopy can be used to visualize φCbK attachment to phage receptors localized at the flagellated pole of swarmer cells (Fig. S3 and Materials and Methods). We investigated whether mutations in *cpaL* and *cpaM* affect the efficiency of φCbK attachment to flagellar filaments and binding to polar localized phage receptors on the cell surface. In comparison to wild-type cells, we observed that Tn*5* insertions within *cpaL* exhibit significant reduction in φCbK attachment to flagella (46.7 ± 14.4%). Fewer phage particles are found to localize in close proximity to the flagellated cell pole of swarmer cells (14.4 ± 6.5%). In contrast, Tn*5* insertions within *cpaM* did not significantly reduce phage attachment to the flagellum (83.7 ± 9.4%) or to the pole of swarmer cells (59.6 ± 6.0%). Similar values were obtained for a Δ*cpaC* mutant that exhibits (81.8 ± 12.0%) flagellar attachment and φCbK phages in close vicinity to the cell surface (53.6 ± 10.1%). In order to successfully infect *Caulobacter* cells, φCbK has to bind with its tail filament to surface receptors located on pili portals (Fig. 1b). In the absence of receptors, although being transported by the flagellar rotation to the cell surface, φCbK fails to efficiently bind with its tail filament to the cell surface. Electron microscopy imaging of a Δ*cpaC* mutant in the presence of phage showed that φCbK phage fails to make contact with the cell surface and most phage tails are pointing outwards (Fig. S3b). Similar results were obtained for a Tn*5* insertion within *cpaM* (Fig. S3b). This suggests that *cpaM* directly or indirectly affects biogenesis of functional pili portals in a similar manner as observed for a Δ*cpaC* deletion. In contrast, although strongly reduced in numbers, we occasionally observed phage particles localized with their tail filament attached to polar sites on swarmer cells of a *cpaL* mutants (Fig. S3b). This suggests that, in a *cpaL*::Tn*5* mutant functional pili, portals are still present but dramatically reduced in numbers. Therefore, it is likely that *cpaL*, rather than being a structural component of the pili secretion apparatus, regulates efficiency of assembly or adopts a regulatory function in pili biogenesis.

Conclusion

Traditional genetic approaches use stringent selective conditions to assign genes to a particular phenotype and map mutants in a one-by-one approach. While such conventional approaches are capable of identifying individual core genes, they usually fail to identify outer circle components of the gene network. However, in a systems-genetics view, components at the edges of gene networks are of equal importance for understanding a particular trait, as these components often contain regulatory functions and provide integration points with other cellular modules.

Here, we have demonstrated how quantitative selection analysis coupled to transposon sequencing (QS-TnSeq) can be employed as a powerful systems genetic tool to examine thousands of genes simultaneously and dissecting their contribution to a particular selectable phenotype. In combination with hierarchical clustering, QS-TnSeq can be used to illuminate gene interaction networks and dissect their submodules and composition. In the case of φCbK infection, it is important to note that not all genes required for φCbK infection can be mapped with our QS-TnSeq approach. Many of the cellular functions hijacked upon bacteriophage infection are encoded within essential core genes of the host cell. Since phages do not encode many cellular core functions, they critically depend on a fully functional metabolism and an intact transcription/translation system of their host. In case of φCbK infection, this includes the complete set of the 480 recently mapped essential *Caulobacter* ORFs [7]. In this context, it is interesting that not only φCbK replication but also its initial attachment and infection mechanisms might use components that, albeit not being essential under laboratory conditions, are of fundamental importance for cell cycle propagation of *Caulobacter* under its natural ecological niche.

In terms of *Caulobacter* biology, our QS-TnSeq results demonstrate the complex interplay between φCbK and the genetic network that drives attachment, infection and replication within its host. Surprisingly, we found that disruptions of only a small subset of 7 flagellar hook-associated genes, corresponding to 8.1% of all 86 annotated motility genes, significantly affect φCbK infectivity whereas disruptions within the remaining 79 motility genes did not impair φCbK susceptibility significantly. Therefore, despite being important for mediating initial phage attachment, the presence of functional flagella does not seem to be a prerequisite for φCbK subsequent infectivity. Our QS-TnSeq results rather suggest that the flagellar hook genes *flgE*, *flgD*, *fliK*, *flbD* and *flgL* identified as contributors of φCbK infectivity constitute components of a signaling mechanism that couples the temporal completion of flagellar hook biogenesis with initiation of biogenesis of the pili secretion apparatus. On the level of φCbK surface receptors, our study confirmed earlier findings that suggested involvement of the pili secretion portals as the entry sites of φCbK DNA. In addition to known structural components of the *Caulobacter* pilus assembly cluster, we assigned several genes that likely represent novel components of the pilus secretion machinery through our QS-TnSeq analysis. Among them are *cpaG* and *cpaH* located downstream of *cpaF* and *cpaI*, *cpaJ* and *cpaK*, which presumably encode for prepilins that facilitate initiation of pilin filament assembly. Two additional genes, *cpaL* and *cpaM*, located outside the *Caulobacter* pili cluster, represent novel structural or regulatory components of the *Caulobacter* pilus assembly system. It will be interesting to investigate how these newly identified pili genes regulate polar organelle development and how they affect the protein localization network that coordinates the dynamic subcellular positioning of the cell fate determinants, DivJ and PleC, and the pili protein CpaE in *Caulobacter*. In sum, QS-TnSeq represents a generally applicable high-resolution approach to map quantitative genotype–phenotype interactions on the genomic scale in a broad class of bacterial species. Beyond phage infection, QS-TnSeq will be of particular interest for systems-genetics analysis of complex bacterial traits including antibiotic persistence phenotypes.

Materials and Methods

Supplementary methods includes descriptions of (i) bacterial strains, growth conditions, bacteriophage propagation and imaging; (ii) transposon construction, mutagenesis and construction of complementation plasmids; (iii) quantitative phage selection assay, DNA library preparation and sequencing; (iv) sequence analysis; (v) data analysis and processing; and (vi) isolation and mapping of individual phage resistance mutants.

Acknowledgements

We would like to thank Catharine Aquino from the Zürich Functional Genomics Center for performing next-generation sequencing, Maya Margrit Günthert from the Scientific Center for Optical and Electron Microscopy (ScopeM), Eidgenössische Technische Hochschule Zürich for electron microscopy and Mirjam Christen for critical reading of the manuscript. This work was supported by the Eidgenössische Technische Hochschule Zürich.

Author Contributions: M.C. and B.C. designed the research. C.B., Y.B., D.C., C.F., F.T., M.C. and B.C. performed the experiments and analysis; C.B., L.D.M., M.C. and B.C. performed the TnSeq analysis; and M.C. and B.C. wrote the manuscript.

Competing Financial Interests: The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2015.11.018.

Received 22 September 2015;
Received in revised form 16 November 2015;
Accepted 17 November 2015
Available online 22 November 2015

Keywords:
transposon sequencing;
systems biology;
type IV pili biogenesis;
bacteriophage;
φCbK;
Caulobacter crescentus;

Abbreviations used:
MOI, multiplicity of infection; PMF, proton motive force.

References

[1] N. Agabian-Keshishian, L. Shapiro, Stalked bacteria: Properties of deoxriybonucleic acid bacteriophage phiCbK, J. Virol. 5 (1970) 795–800.
[2] R.C. Johnson, B. Ely, Analysis of nonmotile mutants of the dimorphic bacterium *Caulobacter crescentus*, J. Bacteriol. 137 (1979) 627–634.
[3] A. Fukuda, M. Asada, S. Koyasu, H. Yoshida, K. Yaginuma, Y. Okada, Regulation of polar morphogenesis in *Caulobacter crescentus*, J. Bacteriol. 145 (1981) 559–572.
[4] M.L. Lawler, D.E. Larson, A.J. Hinz, D. Klein, Y.V. Brun, Dissection of functional domains of the polar localization factor PodJ in *Caulobacter crescentus*, Mol. Microbiol. 59 (2006) 301–316.

[5] J.S. Weitz, H. Hartman, S.A. Levin, Coevolutionary arms races between bacteria and bacteriophage, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 9535–9540.
[6] R.C. Guerrero-Ferreira, P.H. Viollier, B. Ely, J.S. Poindexter, M. Georgieva, G.J. Jensen, et al., Alternative mechanism for bacteriophage adsorption to the motile bacterium *Caulobacter crescentus*, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 9963–9968.
[7] B. Christen, E. Abeliuk, J.M. Collier, V.S. Kalogeraki, B. Passarelli, J.A. Coller, et al., The essential genome of a bacterium, Mol. Syst. Biol. 7 (2011) 528-528.
[8] M. Tomich, P.J. Planet, D.H. Figurski, The tad locus: Postcards from the widespread colonization island, Nat. Rev. Microbiol. 5 (2007) 363–375.
[9] C.A. Whittaker, R.O. Hynes, Distribution and evolution of von Willebrand/integrin A domains: Widely dispersed domains with roles in cell adhesion and elsewhere, Mol. Biol. Cell 13 (2002) 3369–3387.
[10] P. Youderian, N. Burke, D.J. White, P.L. Hartzell, Identification of genes required for adventurous gliding motility in *Myxococcus xanthus* with the transposable element mariner, Mol. Microbiol. 49 (2003) 555–570.
[11] F.F.V. Chevance, K.T. Hughes, Coordinating assembly of a bacterial macromolecular machine, Nat. Rev. Microbiol. 6 (2008) 455–465.
[12] J.M. Sommer, A. Newton, Sequential regulation of developmental events during polar morphogenesis in *Caulobacter crescentus*: Assembly of pili on swarmer cells requires cell separation, J. Bacteriol. 170 (1988) 409–415.
[13] G. Leclerc, S.P. Wang, B. Ely, A new class of *Caulobacter crescentus* flagellar genes, J. Bacteriol. 180 (1998) 5010–5019.
[14] E.K. Mangan, J. Malakooti, A. Caballero, P. Anderson, B. Ely, J.W. Gober, FlbT couples flagellum assembly to gene expression in *Caulobacter crescentus*, J. Bacteriol. 181 (1999) 6160–6170.
[15] E.A. O'Neill, R.A. Bender, Cell-cycle-dependent polar morphogenesis in *Caulobacter crescentus*: Roles of phospholipid, DNA, and protein syntheses, J. Bacteriol. 171 (1989) 4814–4820.
[16] C. Cheng, M.J. Wakefield, J. Yang, M. Tauschek, R.M. Robins-Browne, Genome-wide analysis of the Pho regulon in a *pstCA* mutant of *Citrobacter rodentium*, PLoS ONE 7 (2012) e50682.
[17] M. Gonin, E.M. Quardokus, D. O'Donnol, J. Maddock, Y.V. Brun, Regulation of stalk elongation by phosphate in *Caulobacter crescentus*, J. Bacteriol. 182 (2000) 337–347.
[18] Y. Minato, S.R. Fassio, J.S. Kirkwood, P. Halang, M.J. Quinn, W.J. Faulkner, et al., Roles of the sodium-translocating NADH:quinone oxidoreductase (Na^+-NQR) on *Vibrio cholerae* metabolism, motility and osmotic stress resistance, PLoS ONE 9 (2014) e97083.
[19] J.N. Feng, M. Russel, P. Model, A permeabilized cell system that assembles filamentous bacteriophage, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 4068–4073.
[20] B. Christen, M.J. Fero, N.J. Hillson, G. Bowman, S.-H. Hong, L. Shapiro, et al., High-throughput identification of protein localization dependency networks, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 4681–4686.
[21] Q. Xu, B. Christen, H.-J. Chiu, L. Jaroszewski, H.E. Klock, M.W. Knuth, et al., Structure of the pilus assembly protein TadZ from *Eubacterium rectale*: Implications for polar localization, Mol. Microbiol. 83 (2012) 712–727.
[22] A.O. Gaca, C. Colomer-Winter, J.A. Lemos, Many means to a common end: The intricacies of (p)ppGpp metabolism and its control of bacterial homeostasis, J. Bacteriol. 197 (2015) 1146–1156.
[23] D. Gonzalez, J. Collier, Effects of (p)ppGpp on the progression of the cell cycle of *Caulobacter crescentus*, J. Bacteriol. 196 (2014) 2514–2525.
[24] P.H. Viollier, N. Sternheim, L. Shapiro, A dynamically localized histidine kinase controls the asymmetric distribution of polar pili proteins, EMBO J. 21 (2002) 4420–4428.
[25] A. Marchler-Bauer, M.K. Derbyshire, N.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, et al., CDD: NCBI's conserved domain database, Nucleic Acids Res. 43 (2015) D222–D226.