The Dissertation Committee for Brian Anthony Renda Certifies that this is the approved version of the following dissertation:

Laboratory evolution and natural transformation in *Acinetobacter baylyi* ADP1

Committee:

Jeffrey Barrick, Supervisor

James Bull

Andrew Ellington

Hal Alper

Bryan Davies

Laboratory evolution and natural transformation in *Acinetobacter baylyi* ADP1

by

Brian Anthony Renda, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2016

Acknowledgements

My graduate career has been blessed by the influence of a number of individuals who have helped me become the scientist and person I am today.

First, I would like to thank my graduate advisor Dr. Jeffrey Barrick for the transformative opportunity he provided by accepting me as a founding member of his research group at UT Austin. Throughout the development of my research project, he provided a continuous and steady stream of support and mentorship. The combination of a wholehearted enthusiasm for pursuing the fantastic and a strict adherence to good scientific practice that Jeff seeks to instill in his research group has been a fantastic environment to develop in and has provided the basis for (hopefully) continuing long term success. I would also like to thank committee members Drs. Jim Bull, Andy Ellington, Hal Alper, and Bryan Davies for their support and advice over the course of my research project.

I would also like to thank the various graduate students, post-docs, and undergraduates of the Barrick lab and MBS program for the hands on help, enlightening discussion, and enthusiastic distraction they have provided throughout the years.

Lastly and most of all, I would like to thank my wonderful wife Varsha, my loving daughter Anjali, and the rest of my family for their support and patience throughout my graduate studies and the writing of this dissertation.

Laboratory evolution and natural transformation in *Acinetobacter baylyi* ADP1

Brian Anthony Renda, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Jeffrey Barrick

The bacterium *Acinetobacter baylyi* ADP1 has been proposed as a next-generation chassis for genome engineering due to its high natural transformation rates and metabolic proficiency. Development of ADP1 for biotechnological use however is inhibited by the evolutionary instability of the competence, the large homology requirement for the efficient transformation of DNA into the ADP1 genome, and incomplete knowledge regarding the composition and regulation of the protein complex involved in DNA uptake.

Chapters 1 and 2 describe an investigation of the evolutionary stability of competence. We conducted a 1000 generation evolution experiment with ADP1 and characterized the genetic basis of competence loss. We found that mutations caused by the mobile genetic element IS*1236* mediated loss of were a driving force for genetic instability in the strain. We also found that a previously uncharacterized filamentous phage (CRAϕ) emerged from the ADP1 genome. CRAϕ appears to infect via the ADP1 competence machinery – providing a potential explanation for the fitness benefit of competence loss.

Chapter 3 describes studies aimed at improving transformation rates and better characterizing the ADP1 transformation system. We examined how perturbing factors that affect the fate of internalized DNA effects transformation rates at different homology lengths. We found that some factors, such as exonuclease activity and recombinase activity, limited transformation rates at certain homology lengths and DNA concentrations. Other factors, such as adding random DNA to the ends of fragments to buffer against exonucleases and altering the abundance of a single-stranded DNA protecting protein, did not.

Chapter 4 describes an effort to better characterize the machinery required for competence in ADP1. We coupled transformation-mediated selection of an ADP1 transposon library with Tn-seq to identify genes required for competence. Through this method, we were able to positively identify two additional competence genes (*pilR* and *ACIAD3188*) and other candidates.

Together, the work presented in this dissertation provides insights into the evolutionary stability and molecular workings of ADP1 competence. These insights represent progress toward further improving ADP1 for use as a versatile genome and metabolic engineering platform.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

NATURALLY TRANSFORMABLE BACTERIA – A FAST LANE TO GENETICALLY ENGINEERING DIVERSE SPECIES OF BACTERIA

In bacteriology, the term 'transformation' refers to the transfer of genetic information by naked DNA into cells. First described by the conversion of non-virulent *Streptococcus pneumoniae* to the virulent form by means of heat-killed virulent cells in Griffith's classic experiment(1), the term transformation itself turned out to be quite prophetic. Our ability to manipulate bacterial genetics through the introduction of extrachromosomal genetic elements or direct genome manipulation has transformed biology from a descriptive science to one where application driven technology development is possible.

Since its initial isolation in 1884 and its development as a model organism throughout the 20th century, *Escherichia coli* has been a holistic platform for studying fundamental biological processes, investigating pathogenicity, and developing biotechnology (2). The primary techniques that facilitates such studies are those that allow for the transformation of DNA into the bacteria (3, 4). Transformation into *E. coli* has allowed for both the introduction of plasmids as well as direct manipulation of the bacterial genome itself (5–7). Such techniques have been instrumental in developing *E. coli* into the biotechnology platform it is today, where it is responsible for producing a significant portion of recombinant therapeutic proteins (8) as well as biofuels and other industrially relevant chemicals (9).

Despite *E. coli*'s immense potential as a platform for genetic engineering, there are situations where using other non-model bacterial species can be beneficial (10). In industrial microbial biotechnology, it is common practice to utilize a diverse array of

microbial species that are particularly well suited for performing specific tasks – usually production of specific products at high levels (11). Genetic manipulation of non-model strains is often required for optimizing the productivity of those systems through biological engineering. This means that in order to fully realize the potential of non-model systems, having tools to introduce genetic material and engineer their genomes is vital. Luckily, there is a phylogenically diverse set of bacterial species in which the introduction of genetic material and manipulation of the genome are easy and part of the natural physiology of the species: naturally transformable bacteria. Coupled with beneficial secondary strain characteristics in a number of the species, naturally transformable bacteria represent a wide array of potential engineering platforms for building biological devices.

Naturally transformable bacterial species, also known as naturally competent bacteria, can uptake exogenous DNA and integrate it in a heritable fashion without the need for artificial chemical treatment or electroporation (12). Natural transformation is facilitated by a multi-protein transmembrane complex henceforth referred to as the competence apparatus. The natural transformation phenotype is a broadly distributed across bacterial species – it has so far been found in 8 different bacterial phyla and in more than 60 individual species (13). These systems can transform both linear DNA and plasmids, and transformed DNA can be directly integrated into the bacterial genome through homologous recombination. In this way, generating knockouts and knock-ins of genes is often just a matter of treating the bacteria with homology-containing DNA under competence inducing conditions and then selecting for successful transformants.

A number of naturally transformable species have already been coopted for biotechnological use due to the combination of their high genetic tractability along with other beneficial strain characteristics. Two species that have been of particular interest to

biotechnology are *Bacillus subtilis* and *Deinococcus radiodurans*. *B. subtilis*, a Gram-positive naturally transformable species, has risen to be a major workhorse in producing recombinant proteins due to its robust protein secretion system and ease of genetic manipulation (14, 15). The well characterized transformation system of *B. subtilis* was also used in a genome minimization project, in which 8% of the genome was successfully removed (16). Leveraging the use of competence to deliver DNA into cells, many genetic tools built for *B. subtilis* are made to integrate into the genome rather than being maintained on plasmids (17).

Natural transformability of the radiation and desiccation-resistant bacterium *D. radiodurans* has been used to engineer it for metal remediation and organopollutant degradation in radioactive environments (18, 19). In addition to its ability to transform large genetic constructs, the high resistance to physiochemical stress and ability to digest certain polysaccharides has resulted in Deinococcus species being used as the core chassis organism at companies such as Deinove (Grabels, France) for the production of chemicals (20). Other examples of naturally transformable bacteria that have been functionalized for biotechnological use include *Streptococcus thermophilus* for food production (21) and *Synechocystis* sp. for hydrocarbon production (22). As the number of characterized microbial species grows, it is likely that more naturally transformable bacteria will be found and that some of these will have other desirable properties, increasing the potential impact of natural transformation on pushing forward biological engineering.

ACINETOBACTER BAYLYI ADP1 IS AN EXAMPLE OF AN IDEAL NATURALLY TRANSFORMABLE BACTERIUM FOR BIOTECH APPLICATIONS

Acinetobacter baylyi ADP1 is a Gram negative naturally transformable species that is not currently used in industry but is potentially valuable for diverse applications. Proposed as an ideal model organism for biological engineering (23–25), ADP1 lacks pathogenicity, is metabolic versatile, and has particularly high natural transformability.

History and uses of *Acinetobacter baylyi* ADP1

Prior to its taxonomic reclassification in 2006, *Acinetobacter baylyi* ADP1 was alternatively known as both *Acinetobacter calcoaceticus* BD413 (26) or *Acinetobacter sp.* ADP1 (27). This strain was initially isolated as *Acinetobacter calcoaceticus* BD4 (28) from soil on the basis of its ability to grow on 2,3-butanediol as a sole carbon source (29). ADP1 is a microencapsulated and subsequently non-aggregating variant of BD4 that was isolated via random mutagenesis. A defining characteristic of *A. baylyi* ADP1, its natural transformability, was discovered when mixing encapsulated and unencapsulated mutants of BD4 led to reversion of unencapsulated mutants to an encapsulated form through genetic recombination (28). Since the discovery of transformation in ADP1, efficient transformation has been demonstrated in a handful of other *A. baylyi* strains (30) as well as more broadly across the Acinetobacter genus (30, 31). Interestingly, the conservation of competence-related genes across this genus (32) is much wider than demonstrated transformability. This may indicate that other genus members may be transformable but under unknown conditions (33) or that transformation systems in most Acinetobacter strains have recently evolved to be non-functional.

Since its discovery, much of the research concerning ADP1 has generally centered around its metabolic versatility (25) and natural transformation (discussed

below) (34). There has also been a significant push to functionalize ADP1 to solve specific problems. As a soil bacterium, one use ADP1 is particularly well suited for is as an environmental biosensor. Accordingly, ADP1 has been used to detect degradation of the pollutant naphthalene via salicylate (35), DNA damage inducing compounds (36), and to sense toluene and xylene (37). Leveraging ADP1's natural transformation, the strain has also been developed to detect in situ horizontal gene transfer (38) and to assay for the presence of transgenic plant DNA (39).

In addition to biosensor development, there has been a small but concerted effort to use ADP1 for producing chemicals. The primary focus of this has been the production of wax esters which some Acinetobacters accumulate in large amounts intracellularly (40, 41). Metabolic engineering in ADP1 has resulted in improved wax ester production (42–44), although use of the strains by industry in this capacity yet to be reported. Further efforts have been made to facilitate future metabolic engineering in ADP1, including adding pyruvate kinase (*pykF*), which ADP1 naturally lacks, to enable robust growth on glucose and gluconate (45), and the construction of a global metabolic model in ADP1 to facilitate flux balance analysis calculations (46). Production of a bioemulsifier from ADP1 has also been demonstrated, but has not been subject to production optimization (47). It is likely that further exploration of the biotechnological uses of ADP1 will yield new applications.

Natural transformation in *Acinetobacter baylyi* ADP1

As mentioned previously, naturally transformation allows bacteria to bind, uptake, and integrate extracellular DNA from their environments as part of their normal cellular processes. Generally, this process is facilitated by a transmembrane multi-protein

competence apparatus that is homologous to type IV pili and type II secretion systems (48). A number of natural transformation systems have been studied in depth and have been the subject of exhaustive reviews (12, 49, 50). While not nearly as well studied as other naturally transformable species, a body of work has arisen that gives key insight into the biology of ADP1 natural transformation. Remaining gaps in knowledge of transformation mechanisms in ADP1 can be reasonably guessed at from studies in other Gram-negative strains, as these competence systems seem to be related based on conservation in their protein sequences and structures. Natural transformation can be broken down into three phases: 1) development of the competent state, 2) binding and translocation of extracellular DNA, and 3) processing and integration of transformed DNA.

Development of competence

Natural competence, the physiological state of the bacterium that allows for natural transformation to occur, is subject to highly divergent regulation between different species of naturally transformable bacteria (51). While some species like *Neisseria gonorrhoeae* are constitutively competent (52), in other species competence is triggered by nutritional stress (53), a quorum sensing pheromone (54), environmental inducers (55), or antibiotic induced stress (56). In ADP1, competence is induced upon a nutrient upshift after growth cessation – generally embodied as the beginning of regrowth after a stationary phase culture is transferred into fresh media (57).

How the induction of competence in ADP1 relates to the regulation of competence genes is currently unclear. Only competence genes *comP* and *comB* have had their expression profiles measured in relation to growth phase – both using *lacZ* fusion reporter genes on low copy plasmids. Expression of *comP* is maximal during stationary

phase and drops greatly upon dilution into fresh media, only to increase again as the culture reaches late log and stationary phase (58). Expression of *comB* follows a similar pattern with a marginal increase immediately after dilution into fresh media before rapidly decreasing (59). As with *comP*, *comB* expression increases again as stationary phase is reached. As the expression of these structural pilin genes does not correlate with competence induction, it appears that the competence apparatus is, at least in part, synthesized prior to the maximal induction of competence. It has been proposed that competence induction in ADP1 is indeed independent of protein production as ADP1 is transformable in the presence of chloramphenicol, an inhibitor of protein synthesis (60). The lack of correlation between competence gene expression and competence induction suggests that the assembly and activity of the competence pilus could mirror that of other characterized type IV pilin assembly in response to chemotactic stimulation, where intracellular sensors undergo conformation changes and interact with pilin ATPases to mediate pilus assembly and retraction (61). This model would be consistent with the quick induction of competence upon dilution into fresh media and explain the apparent disconnect between competence induction and competence gene expression.

DNA binding and translocation

After induction of the competent state and assembly of the competence apparatus, DNA binding and translocation can occur. While there has been no mechanistic study of this process in ADP1, some work characterizing specific competence and models informed from related systems outlines the following general model. Extracellular double stranded DNA (dsDNA) is bound to the portion of the competence apparatus presented on ADP1's outer membrane, likely including ComC, a type IV pilin whose location and ability to bind DNA are consistent with a potential role in initial DNA binding and

passing that DNA onto the transformation pilus (62). In ADP1, a transformation pseudopilus consisting of type IV pilin-like proteins ComP (58, 63), ComB (59), ComE (64), and ComF (64), then likely extends through the secretin ComQ in the outer membrane to bind the transforming DNA (34). The four putative core members of the transformation pseudopilus, interestingly, do not appear to contribute equally to the ability to transform DNA in ADP1 – loss of *comB* or *comP* results in complete abolishment of transformability (58, 65), while loss of *comF* or *comE* result in only 1000-fold and 10-fold reductions, respectively (64). This demonstrates that while the presence of all four subunits is required for optimal function, only a subset are absolutely required for pseudopilus function. The competence pseudopilus, either by its own DNA binding capacity or through the action of a shuttling protein ComEA (34), then drags the transforming DNA into the cell's periplasm via retraction of the pseudopilus by depolymerization of the pilus structure. Once inside the periplasm, the dsDNA is loaded into the inner membrane DNA translocator ComA, although other competence genes may also be involved in the process (34). When DNA travels through ComA, which is predicted to form a multimeric inner membrane pore, it is converted into single stranded DNA (ssDNA) and enters the cytosol in the 3′ to 5′ direction (60). The energetics of this process are unclear – studies conducted in other species and reviewed in depth elsewhere (60) suggest that in some species that uptake might be related to the proton motive force while in others intracellular ATP levels are important. While mechanistic studies in ADP1 are lacking, it is possible that pH, ATP levels, or both could provide the energy for DNA uptake. Alternatively, hydrolysis of the non-transforming strand of the transported DNA could provide the energy for uptake for translocation into the cytosol (66).

DNA processing and integration

Once in the cytosol, the now ssDNA has two potential fates: degradation by cellular nucleases or integration into the ADP1 genome. In bacteria, a number of nucleases with single stranded activity have been characterized including ExoI (67), ExoVII (68), SbcCD (69), ExoX (70), and RecBCD (71). Generally these exonucleases function in recombination and DNA repair (72), but theoretically any ssDNA nuclease activity could reduce the pool of cytosolic transformed DNA and change transformation rates. In ADP1, deletion of *recBCD* result in a small but significant increase in transformation frequency, albeit at the cost of lower cell viability (73), suggesting a potential role of the enzyme in degrading transformed DNA. Another exonuclease in ADP1 is RecJ, which is a ssDNA specific 5′→3′ exonuclease that plays a key role in *recBCD*-independent recombination (74). RecJ is highly processive exonuclease and has the ability to degrade ~1000 bases per binding event (75). It has also been shown that single stranded binding protein interacts with RecJ to recruit RecJ to ssDNA for degradation (75, 76). In ADP1, RecJ has been shown to suppress homology-facilitated illegitimate recombination with small amounts of homology and improve transformations with short DNA fragments (77, 78).

In ADP1, SbcCD has also been found to impact homology-facilitated illegitimate recombination and homologous transformation with small amounts of homology, but not impact transformations with large amounts of homology (79). ExoI and ExoVII have no identifiable homologs in ADP1, but a homolog to ExoX with ssDNA nuclease activity has recently been identified (78). Interestingly, loss of *exoX* does not alter transformation by itself. When combined with a *recJ* knockout loss of *exoX* increases transformation frequencies for short DNA products by more than 100-fold, an effect which again goes away at longer DNA homology lengths. Together, these observations demonstrate that in

ADP1, transforming DNA is subject to exonuclease degradation but that limitation can be overcome by simply transforming with longer DNA fragments that maintain sufficient homology to the ADP1 chromosome even when partially degraded.

If the transforming DNA is not degraded by cytosolic nucleases, it is accessible to recombinase activity for integration into the genome. In ADP1, this process is mainly facilitated by RecA, which enables homologous recombination into the ADP1 genome. To aid the process of loading internalized ssDNA onto RecA, many naturally transformable bacteria including ADP1 have a protein that can bind and protect ssDNA and then pass it onto RecA called DprA (80). In some species, having *dprA* is required for competence (81), while loss of the gene in ADP1 results in only ~10-fold reduction (34). With successful nucleation of RecA onto the transformed ssDNA, homologous recombination can proceed through its normal mechanism (82).

PATHS FORWARD TO DEVELOP UNTAPPED POTENTIAL FOR NATURALLY TRANSFORMABLE BACTERIA

Currently, naturally transformable bacteria are greatly underutilized in biotechnological applications for a number of reasons. The better characterized transformation systems reside in pathogenic bacteria. Many transformation systems are also only induced under specific environmental conditions such as the presence of chitin (55), a specific pheromone (83), or cell stress (84). The need for species specific recognition sequences in DNA for it to be bound and internalized by the competence apparatus in a number of species (85, 86) is also particularly problematic for biological engineering due to the sequence constraints it creates. *Acinetobacter baylyi* ADP1 is competent during normal laboratory growth, is not pathogenic, and takes up any DNA

regardless of its sequence. These characteristics make it a particularly useful platform to explore applications of natural transformation in biological engineering.

One application in which natural transformation has great promise is in increasing the evolvability of engineered systems. Transforming naturally competent bacteria with PCR products amplified with error prone polymerases (PCR mutagenesis) has been shown to be an effective tool in probing the function of genes in ADP1 (25). A natural extension of this would be to introduce variation onto a subset of specific genes to select for optimization of beneficial phenotypes. In *V. cholera*, such an approach with a defined pool of variation has been effective for multiplex genome engineering to improve transformation rates in this species (87). This was accomplished by refactoring the promoters and ribosome binding sites for five genes related to competence gene regulation or DNA processing, using cotransformation to create a diverse mutant pool, and then two rounds of serial selection to enrich the population for clones with improved competence. The result of this process was the creation of an individual mutant with ~30-fold increased transformability due to increased expression in *recA* and a regulator of competence apparatus biosynthesis, *tfoX* (87). Beyond the targeted introduction of variation, natural transformation gives a useful tool that is otherwise absent from bacterial engineering: sexual recombination. In situations that require adaptive laboratory evolution to improve strain growth or performance, being able to pass genes between members of an evolving population allows for the shuffling of beneficial alleles and can offer more mobility as an organism navigates fitness landscapes towards improvements (88).

Interestingly, the ADP1 transformation apparatus itself has not yet been subjected to an engineering effort. The system, which binds extracellular DNA as double-stranded DNA and imports it as single-stranded DNA, could have a number of uses after sufficient

progress in the understanding the details of its molecular operation. For example, as transformed DNA is converted into ssDNA and enters the cytoplasm, the presence of a ssDNA-targeted RNA polymerase such as that from bacteriophage N4 (89) along with the appropriate promoter sequence on the transforming DNA could result in the direct expression of transformed DNA without the need for plasmid expression or genome integration. With a complete understanding of the composition of the transformation apparatus, one could port this system into a wide variety of organisms, imparting the utility of natural transformation and circumventing the need for independent development of genome engineering techniques for each species. Indeed, a full understanding of natural transformation could provide valuable tools for use in the next generation of engineered biological machines.

In the following body of work, I present studies investigating *Acinetobacter baylyi* ADP1 and its natural transformation system. Overall, the goal of this has been to improve the utility of the organism in biological engineering by learning more about three major problems in ADP1: the evolutionary instability of competence (Chapters 2 and 3), the large homology requirement for efficient transformation into the ADP1 genome (Chapter 4), and the incomplete characterization of the composition and regulation of ADP1's natural transformation machinery (Chapter 5). As a result of this work, we have a better understanding of how to maintain competence in the ADP1 strain during laboratory culture, have a better framework for understanding how to improve transformation rates, and know more about the genes required for natural competence.

Chapter 2: Genome instability & loss of competence in *Acinetobacter baylyi* ADP1[1]

INTRODUCTION

The large-scale engineering of genome sequences is a growing area of synthetic biology and bioengineering with many applications (90, 91). Projects such as the development of reduced bacterial genomes (92), modifying the genetic code (93), and optimizing metabolic pathways (6) require whole-scale genetic rewriting of genomes in order to accommodate progressively more ambitious goals. The process of making such modifications is a resource intensive undertaking with significant technical, human-resource, and time costs. *Acinetobacter baylyi* ADP1 has been proposed as a model organism for metabolic engineering, synthetic biology, and genome evolution studies because it is naturally transformable (23, 24, 94). During normal growth in the laboratory, ADP1 expresses competence machinery that enables it to efficiently import extracellular DNA without any sequence constraints (95, 96). By designing DNA constructs with homology to the ADP1 genome, targeted deletions or insertions of new genes can be made with high efficiency without the need for artificial transformation techniques or the expression of heterologous genes, as is typically required when engineering the *Escherichia coli* genome (6, 97). Thus, ADP1 could be a useful chassis organism for studies that require large-scale or combinatorial genome modifications.

One known challenge associated with ADP1, however, is that the competence phenotype is unstable when it is cultured in the laboratory (98). In general, genetic

[1] Chapter is reproduced (with minor modifications) from its initial publication:

Renda, B. A., Dasgupta, A., Leon, D. & Barrick, J. E. Genome Instability Mediates the Loss of Key Traits by Acinetobacter baylyi ADP1 during Laboratory Evolution. *J. Bacteriol.* **197,** 872–881 (2015)

Contributions: A.D. performed the mutation accumulation experiment. D.L. performed the light microscopy. J.E.B. assisted in the experimental design and writing of the paper.

reliability is a key trait sought after in chassis organisms; one does not want to design a DNA sequence only to see it be rapidly deleted or mutated when placed in a living organism. In this vein, efforts to engineer *E. coli* variants with more stable genomes have resulted in important benefits in those strains (92, 99). In ADP1, the frequent loss of transformability would be particularly problematic for whole-genome engineering efforts, which may require many steps of selection and regrowth to reprogram multiple genetic loci. Therefore, it would be beneficial to understand the genetic basis of competence loss in ADP1 in order to potentially design strategies to stabilize this highly desirable trait.

As ADP1 has not yet been propagated in the laboratory to the same extent as bacteria such as *Escherichia coli* or *Bacillus subtilis*, it also more generally represents a model for moving a relatively undomesticated microbial species into the laboratory. In working with this kind of organism, a researcher generally will want to ascertain the dominant forces of genome instability at play (100) and eliminate any that are at odds with bioengineering and synthetic biology goals. In the wild, bacterial genomes often accumulate cryptic bacteriophage remnants and selfish genetic elements, such as transposons. These sequences may lead to certain regions of an undomesticated host genome being particularly prone to mutations or to certain types of mutations dominating among those that inactivate engineered DNA sequences deployed in the host.

To determine the genetic basis of competence loss and to profile the stability of the ADP1 genome as a whole, we sequenced the genomes of seven ADP1 strains that were independently evolved for 1,000 generations in an adaptive evolution (AE) experiment in the laboratory. In parallel, we conducted a mutation accumulation (MA) evolution experiment in which replicate populations of ADP1 were propagated for ~7,500 generations of genetic drift. In the AE experiment we expected to see mainly beneficial mutations become prominent within populations, whereas in the MA

experiment single-cell bottlenecks between transfers would allow non-lethal mutations to accumulate regardless of their impact on fitness. The AE experiment reproduced the expected decrease in natural transformation in the evolved populations. We also observed the unexpected reevolution of a cellular aggregation phenotype and parallel losses of a large prophage-related region containing several genes that were thought to be essential. The loss of transformability and increased aggregation phenotype were overwhelmingly caused by the activity of IS*1236*, the sole mobile genetic element in the ADP1 genome. These three types of mutations were never or only rarely observed in the MA evolution experiment, supporting the hypothesis that they evolve because they improve fitness in nutrient broth.

MATERIALS & METHODS

Strains and growth conditions

The *Acinetobacter baylyi* ADP1 strain used for this work was the gift of Valérie de Crécy-Lagard (University of Florida). Unless otherwise noted, ADP1 was cultured at 30°C on the Miller formulation of Luria-Bertani (LB) medium (10 g/L NaCl). For tests of gene essentiality, minimal succinate (MS) (101) medium was used. Liquid cultures were growth in 50-mL Erlenmeyer flasks or 18-mm by 150-mm test tubes with orbital shaking at 140 r.p.m. over a 1-inch diameter. Solid LB media included 1.5% (w/v) agar. Where appropriate, media were supplemented with 50 μg/mL kanamycin or 200 μg/mL 3'-azido-2', 3'-dideoxythimidine (AZT). Frozen stocks of strains and evolved populations were stored at -80°C with added glycerol (20% v/v) as a cryoprotectant.

Adaptive evolution experiment

Eight colonies of the ancestral ADP1 strain were inoculated into separate flasks containing 10 mL LB to initiate experimental populations 1-8. These cultures were propagated from a total of 100 days with the serial transfer of 10 µL into a new flask containing 10 mL LB every 24 ± 1 hours of incubation. This 1:1,000 dilution resulted in the transfer of ~10^7 cells from each saturated 24 hour culture and ~ 10 generations of regrowth per day. Samples of each population were saved as frozen glycerol stocks every 50 generations. At each transfer, populations 1 to 4 were supplemented with 100 ng of metagenomic DNA isolated from soil collected at the Brackenridge Field Research Laboratory (Austin, TX) as previously described (102) but scaled to process 80 g of soil at a time. One population with supplemented DNA was discontinued midway through the experiment due to persistent contamination.

Mutation accumulation evolution experiment

Twenty-five different ADP1 ancestor colonies were picked from an LB agar plate to initiate independent lineages. Each line was maintained by picking a colony each day and streaking it out on a fresh agar plate. After growth for 24 h, the new colony that appeared nearest the end of the streaking pattern was selected to continue the line. Every 20 days, the selected colonies for each line also were used to grow liquid cultures that were stored as frozen glycerol stocks. After 300 transfers, we PCR amplified and sequenced a portion of the 16S rRNA gene using universal primers U341F and UA1406R (103), and we compared this sequence to the ADP1 ancestor to rule out contamination. Seven of the lineages became contaminated or were lost due to other experimental errors.

GFP strain construction

A variant of the ADP1 ancestor with chromosomally encoded GFP was constructed using the pIM1463 integration vector (104). Briefly, the P_{T5}-lacO-lacO-*gfp* insert from pIM1522 (104) was PCR amplified with primers to add SfiI restriction enzyme sites to each end. This GFP construct was digested with SfiI and ligated to the pIM1463 plasmid backbone cut with SfiI. After transformation into ADP1, a colony that exhibited constitutive GFP expression was selected.

Fitness assays

Evolved population samples were competed against the GFP-tagged ancestor by first reviving 12 µl or 2 µl, respectively, of glycerol stocks into 5 mL LB cultures in test tubes. After 24 h of growth, these cultures were used to inoculate 12 flasks containing 10 mL LB for each sample. After 24 h, 12 replicate competitions for each population were initiated by mixing 1.7 µL of each evolved population flask with 8.3 µL of each GFP-ADP1 flask in 10 mL LB. Immediately after mixing and after 24 h of growth, dilutions of these cocultures were plated on LB agar. Counts of fluorescent versus nonfluorescent colonies were used to calculate relative fitness as a ratio of Malthusian parameters as previously described (105).

Whole genome sequencing

Clonal isolates from each 1,000-generation population from the adaptive evolution experiment were inoculated into LB and grown overnight. DNA was extracted from these cultures using the PureLink genomic DNA minikit (Life Technologies) and sheared to an average size of 550 bp using a Covaris S2 instrument. Illumina DNA

fragment libraries were prepared using a NEBNext preparation kit and sequenced on a HiSeq 2000 instrument at the University of Texas at Austin Genome Sequencing and Analysis Facility. Overall read depth coverage of the ancestral genome sequence was ~500-fold for each sample. Mutations were identified from the resulting 2x100-bp paired end reads by comparing them to the complete ADP1 genome sequence (106) (GenBank accession no. NC_005966.1) using the breseq computational pipeline (107).

Detecting horizontal gene transfer

To determine whether genes had been acquired from the metagenomic DNA added to some populations, read pairs where one or both reads did not map to the ADP1 reference genome in the breseq analysis were assembled using velvet (version 1.2.07) with a hash length of 51 bases (108). All resulting contigs with a coverage read depth of at least 10 and length of at least 200 bases were queried against the NCBI nonredundant (nr) database using BLASTN (version 2.2.28+) with default parameters (109). All contigs consisting of high quality reads with mixed base compositions matched the ϕX174 genome (which is added as a spike-in control during Illumina sequencing) or most likely were derived from other DNA samples sequenced in the same lane or flow cell (e.g., matches to *Saccharomyces cerevisiae*, mitochondrial genomes, and cloning vectors). The latter likely were due to barcode misclassifications when demultiplexing reads from different samples.

Mutation rate statistics

We used R (version 3.1.0) for all statistical analyses (110). Poisson regression was used to compare the rates at which different types of mutations accumulated in the

sequenced clones from the genome sequencing data. Here, we evaluated the significance of saturated models with per-clone mutation rates versus models with one uniform rate across all clones using chi-squared tests. To compare the rates of genotypic and phenotypic evolution between the adaptive evolution and mutation accumulation experiments for evidence of selection, we evaluated the significance of Poisson models with distinct rates for each type of evolution experiment to models with one uniform rate across all experiments using chi-squared tests. For these analyses, we assumed that there was only one mutation per experimental lineage, when the same clone theoretically could have accumulated multiple mutations affecting transformability or cellular aggregation. Since many mutations were observed in the adaptive evolution experiment and few were observed in the longer mutation accumulation experiment, this assumption is expected to lead to overly conservative estimates of significance

Prophage deletion detection

We performed two PCR assays to test for the prophage deletion in a given ADP1 strain. The first used one primer outside the deleted region and one primer inside to yield an ~800-bp product if the prophage region still was present but no product if it had been deleted. The second assay used two primers located outside the prophage region to amplify across the deleted sequence, so that there was a ~950-bp product if the deletion was present and no product otherwise.

Integration of kanamycin resistance cassette

To generate an ADP1 strain to serve as a source of genomic DNA for transformation assays (AB-KAN), we integrated a kanamycin resistance marker into the nonessential gene *ACIAD0135*. Upstream (coordinates 139843 to 139832) and downstream (coordinates 140099 to 149928) ADP1 genome segments were PCR amplified with primers that added an EcoRI site to the 3′ end of the upstream segment and a BamHI site to the 5′ end of the downstream segment. Tn5 neomycin phosphotransferase (*nptII*) was amplified from the cloning vector pKD13(97) while adding an EcoRI site to its 5′ end and a BamHI site to its 3′ end. These restriction sites were used to ligate the upstream and downstream ADP1 genome segments to *nptII*. The ligation product was amplified by PCR and then transformed into ADP1 by diluting 70 µL of an overnight ADP1 culture into 1 mL of LB and allowing growth for 5 h under standard conditions. A portion of this culture was plated on LB-kanamycin agar. After growth, resistant clones were picked and assessed for the expected insertion by PCR.

Transformation frequency measurements

AB-KAN genomic DNA samples were isolated using the PureLink genomic DNA minikit (Life Technologies) with yields estimated from absorbance using a NanoDrop spectrophotometer. *A. baylyi* cultures were revived by inoculating a frozen glycerol stock (2 µL for clonal isolates or 12 µL for mixed populations) into 5 mL of LB in a test tube and growing them for 18 to 24 h under standard conditions. Then, these cultures were preconditioned by transferring 10 µL of revived cultures into flasks with 10 mL of LB and growing them for 24±1 h. Transformation assays were initiated by combining 1 mL LB, 70 µL of preconditioned culture, and 100 ng of AB-KAN genomic DNA. These cultures were incubated for 18 to 24 h under standard growth conditions. Transformed

cultures were diluted in sterile saline and plated on both nonselective (LB agar) and selective (LB-kanamycin agar) plates and allowed to grow overnight. Transformation frequencies were determined by dividing the number of CFU on selective media by the number of CFU on nonselective media. All transformation assays were carried out in triplicate. For the adaptive evolution experiment, a matched set of ADP1 ancestor measurements was made alongside each sequenced clone for statistical comparisons. For the mutation accumulation experiment, all endpoint clones were tested at the same time as one set of ADP1 ancestor measurements for comparison. We computed all statistics for transformation frequencies using log-transformed values.

Mutant construction

The ADP1 ancestor strain was modified using a dual-selection cassette as described previously (24). PCR products containing mutant *pgi* and *per* alleles were amplified from clone 6a with an additional 1 kb of flanking homology on either side and used to transform *ACIAD0091* and *ACIAD0101* knockout strains, respectively, from the ADP1 single-gene knockout collection (101). These transformations were plated on LB-AZT agar to select for the mutant allele replacing the Kan^R-*tdk* cassette present at the target location in the knockout library clone. Allelic exchange was verified by PCR and screening transformants on LB-Kan agar to test for the loss of the cassette. The double *pgi per* mutant was constructed by first taking the strain reconstructed with the evolved *per* allele and transforming it with a PCR product containing the integrated Kan^R-*tdk* cassette and 1 kb of flanking genome homology amplified from the *pgi* (*ACIAD0101*) knockout. The evolved *pgi* allele from clone 6a then was used to replace the cassette in this strain as described above. The Δ*pilB* strain, used for electron microscopy, was

created by inserting the Kan^R-*tdk* cassette within *pilB* and then replacing it with a PCR product to create the same in-frame deletion present in the ADP1 knockout collection (101).

Emulsification assay

The secretion of bioemulsifier from cultures was measured as previously described (111). Briefly, 500 μL of supernatant from preconditioned overnight cultures grown in LB was combined with 1 mL of TM buffer (20 mM Tris-HCl, 10 mM MgSO4, pH 7.0) and 20 μL of a 1:1 (v/v) mixture of n-hexadecane and 2-methylnaphthalene in a 10 mL glass test tube. Each sample was vortexed vigorously at room temperature for 30 min. The mixture then was transferred to a glass cuvette, and the optical density at 600 nm (OD600) was measured using a spectrophotometer after the mixture was left to stand for 30 s after vortexing ceased. One emulsification unit equals an OD600 of 0.1.

Microscopy

For light microscopy, we spotted 1 μL of an overnight culture of each strain on a glass slide and imaged it using a Zeiss Axiovert 200M inverted microscope. Images were acquired using differential interference contrast (DIC) with a 63X objective under the transmitted-light bright-field setting using AxioVision SE64 software (release 4.9.1). For electron microscopy, we washed overnight cultures twice in sterile H_2O and then performed negative staining using a 1 or 2% uranyl acetate solution. We then dried the samples onto Formvar-coated copper grids and visualized them with an FEI Tecnai transmission electron microscope.

RESULTS

Acinetobacter baylyi ADP1 evolution experiments

We conducted a 1,000-generation adaptive evolution (AE) experiment in which 7 populations of *Acinetobacter baylyi* ADP1 founded from genetically identical clones were passaged daily in LB nutrient broth either with (populations 2 to 4) or without (populations 5 to 8) metagenomic DNA from forest soil. As expected, the fitness of these populations relative to the ancestral ADP1 strain, as measured in coculture competition assays, improved over the course of the AE experiment (**Figure 2.1**). We isolated a single endpoint clone from each of the 1,000-generation populations for further characterization and sequenced its genome. We did not find any evidence of horizontal gene transfer into these strains (see Materials and Methods) or observe any differences in evolutionary outcomes between populations evolved with or without added soil DNA. Therefore, we do not distinguish between these two types of AE populations in further reporting our results.

In parallel, we performed a 300-day mutation accumulation (MA) evolution experiment involving 18 lineages founded from the same ADP1 ancestor. These lines were propagated on LB agar by choosing a random colony each day, streaking it out on a fresh agar plate, and allowing 24 h for new colonies to grow. Each cycle of regrowth from a single cell to a colony requires ~25 cell divisions (112). The daily single-cell bottlenecks result in greatly reduced selection, such that the rates of genotypic and phenotypic evolution approximate those resulting from unfiltered spontaneous mutations (113). Thus, these endpoint MA clones, which have gone through ~7,500 total generations of growth, were used to determine to what extent mutations in the AE

experiment could be explained by mutational hotspots versus selection for increased fitness.

Genome evolution

After accounting for 32 discrepancies between the ADP1 reference genome sequence(106) and our ancestral ADP1 strain (**Table 2.1**), we found between 6 and 17 mutations in each of the seven independently evolved clones selected for sequencing from the AE experiment (**Figure 2.2A**; also see **Table 2.2**). Similar numbers of new base substitutions and short indel mutations were observed in each clone, and one had a 117 bp duplication. All seven clones had identical 49 kb deletions encompassing a putative prophage (106). We confirmed that the genomic region removed by this large prophage deletion (LPD) was present in the ancestral strain by PCR. In addition to the LPD deletion, at least one other multi-kilobase deletion was observed in most genomes.

The greatest variation in the number of mutations between the sequenced clones was in the activity of IS*1236*, the only native transposable element in ADP1 (106), both in terms of new copies inserting into genes and through its role in facilitating multigene deletions (**Figure 2.2B**). IS*1236* was responsible for 41% of mutations across all sequenced genomes, but the number in individual clones ranged from 0 to 11. In fact, the rate at which new IS element-related mutations accumulated was significantly different among the sequenced clones ($P = 0.00012$ by chi-squared test comparing Poisson rate models), while the rate of all non-IS mutations considered together did not vary ($P = 0.96$). In addition to the LPD, we observed that a number of the same genetic loci were affected by parallel mutations in multiple sequenced clones. *ACIAD3148*, a hypothetical gene of unknown function that is conserved in other Acinetobacter species, incurred

mutations in all 7 populations (**Figure 2.3A**). Since these included two frameshifts and an internal deletion mediated by a new IS*1236* insertion, it is likely that disrupting this gene is highly favorable under the conditions of the adaptive evolution experiment. The *clpA* gene, which encodes an ATPase and specificity subunit for the Clp protease (114), was affected by mutations in 5 of the 7 clones (**Figure 2.3B**). These mutations included a synonymous base substitution in the *clpS* gene that could affect *clpA* gene expression, because *clpA* appears to be transcribed from promoters located in this upstream gene in *A. baylyi* ADP1, as it is in *E. coli* (115). Three other mutations frame shift or duplicate amino acids at the extreme C terminus of the ClpA protein sequence. Interestingly, genomes with mutations affecting *clpA* accumulated IS*1236*-related mutations at a greater rate than those without (P = 0.00054 by one-tailed Poisson test). As discussed below, it is possible that these mutations alter ClpA activity in a way that increases transposase activity. Two of the six IS*1236* copies in the ancestral genome flank the Tn*5613* composite transposon (116). The intervening sequence within this transposon was deleted in four of the sequenced clones, apparently due to homologous recombination between the two flanking IS copies. Some regions deleted by IS*1236*-mediated events also were conserved between multiple populations, including overlapping ~16 kb and ~24 kb deletions in clones 3a and 8a, respectively. Interestingly, clone 6a also contained two mutations, in the *per* and *pgi* genes, within the common deleted region. The only other ADP1 gene affected by a mutation in more than two of the sequenced genomes was *ACIAD2521*. The exact same base substitution was present 127 bp upstream of the start codon of this putative gene of unknown function in four of the sequenced clones, indicating that it was likely present as a polymorphism in the ADP1 culture used to initiate the AE experiment.

IS*1236*-mediated loss of competence

The frequency at which a genetic marker was transformed into the 1,000-generation ADP1 populations from the AE experiment was reduced 10- to 100-fold compared to that of the ancestral strain (**Figure 2.4A**). While all evolved populations had reduced transformability, only three of the sequenced clones (3a, 6a, and 8a) showed significant decreases in transformation frequency ($P < 0.05$ by Bonferroni-corrected one-tailed Welch's *t* test). Only in clones 3a and 8a was this decrease greater than 10-fold. Loss of transformability in clones 3a and 8a can be genetically attributed to insertions of new IS*1236* copies into the genes *pilB* and *comC*, respectively. The observed 100-fold decrease in *pilB* mutants and complete loss of transformation in *comC* mutants are consistent with previous reports indicating that PilB is important for DNA uptake(34) and that the loss of ComC renders a strain completely noncompetent (62). Although the PilB protein is related to type IV pilus components(34), knocking out this gene in the ancestor did not affect piliation (**Figure 2.5**), confirming that it seems to have a function that is specific to DNA uptake. Clone 6a, the isolate with only a 6-fold-reduced transformation frequency, does not have mutations in genes known to affect DNA uptake or recombination (**Table 2.2**).

Measurements of the transformability of additional clones isolated from populations 3 and 8 at the 1,000-generation endpoint of the AE experiment revealed the existence of two distinct sub- populations: fully competent and reduced or noncompetent cells (**Figure 2.4B**). Assaying these clones for *pilB* and *comC* IS*1236* insertions by PCR showed a clear correspondence between the presence of the *pilB* or *comC* mutation and reduced transformation for an isolate. Both populations had one exception to this relationship, in which the mutation from the sequenced clone was not present in an isolate that was less transformable, suggesting that other loss-of-competence mutations

also are present in each evolved population. Finally, we profiled the transformability of clones in population 3 at 200-generation intervals throughout the AE experiment (**Figure 2.4B**). Low-transformation-frequency clones first appeared in this population by 600 generations and became a majority of the population by 800 generations. By 1,000 generations, the less competent subpopulation still had not completely taken over in either of the two populations that we profiled. We expect that the other five populations are similarly composed of mixtures of fully competent cells and cells with reduced competence, but that the clones selected for sequencing happened to be from the fully competent subpopulations in these cases.

If mutations reducing transformability occur at a sufficient rate, it is possible that the loss of this trait would be expected on the time scale of the AE experiment, even if it conferred no selective advantage. To evaluate this hypothesis, we assayed the transformability of the endpoint clones from the 18 MA lines that evolved under conditions of relaxed selection for 7,500 generations. None of the MA clones had a transformation frequency that was reduced, with certainty, more than 2-fold relative to that of the ADP1 ancestor ($P < 0.05$ by Bonferroni-corrected one-tailed Welch's *t* test) (**Figure 2.6**). The rate at which loss-of-transformability mutations appeared in the sequenced AE clones is significantly accelerated relative to this non- selective MA rate ($P = 0.00002$ by chi-squared test comparing Poisson rate models). Therefore, we conclude that ADP1 competence likely is unstable because cells that lose this trait, at least via the mutations we observed, obtain a competitive fitness advantage. It is also possible that the spread of loss-of-competence mutations through a population by horizontal gene transfer further accelerates the loss of transformability in the AE experiment (98).

Evolution of cellular aggregation

We observed that cultures of a number of our sequenced clones from the AE experiment sedimented in test tubes more readily than the ancestral ADP1 strain. This phenotype correlated with distinct patterns of cellular aggregation observed by light microscopy (**Figure 2.7A**). Clones 2a, 4a, 5a, and 7a behaved similarly to the ADP1 ancestor: a majority of cells were attached to no more than one other cell. In contrast, most cells were part of multicellular clumps in clones 3a, 6a, and 8a. Mutations in two genes, those for perosamine synthetase (*per*; *ACIAD0095*) and glucose-6-phosphate isomerase (*pgi*; *ACIAD0101*), were present exclusively in the three aggregating clones. The *per* and *pgi* genes are lost due to multi kilobase IS*1236*-mediated deletions in clones 3a and 8a. In clone 6a, *pgi* is interrupted by an IS*1236* insertion, and *per* contains a single-base deletion leading to a frame shift (**Figure 2.7B**). To determine whether these loss-of-function mutations lead to increased cellular aggregation, we reconstructed the *per* and *pgi* mutations from clone 6a, both independently and together, in the ancestral ADP1 back- ground. We found that each single mutant and the double mutant aggregated (**Figure 2.7C**), indicating that the loss of either gene leads to this phenotype.

Both *per* and *pgi* have been implicated in the biosynthetic pathway of a secreted emulsifier in the closely related species *Acinetobacter lwoffii* RAG-1 (117, 118). A number of other Acinetobacter strains, including ADP1, also can synthesize a bioemulsifier consisting of a protein and carbohydrate component that likely allows them to better utilize fatty acids and hydrocarbons as carbon sources (47, 111, 119). In the emulsifier biosynthetic pathway in RAG-1, *weeJ*, a homolog of the ADP1 per gene, is thought to be involved in the synthesis of one of the three hexose subunits used to construct the polysaccharide component of the emulsifier, and *pgi* has a key role in generating glucose-6-phosphate to feed into this pathway. Therefore, we assayed for

bioemulsifier production in the 6a clone and reconstructed *per* and *pgi* mutants (**Figure 2.7D**). We found that either mutation on its own caused a significant decrease in the emulsification activity of a culture ($P < 0.05$ by one-tailed Welch's *t* test). As was the case for the aggregation phenotype, the level of bioemulsifier production in the double mutant appears to be no different from that measured in either single mutant ($P > 0.05$ by two-tailed Welch's *t* test).

We again examined the MA lines to determine if there was evidence that evolution of cellular aggregation was accelerated in the AE experiment. We found that 4/18 endpoint MA clones exhibited similar aggregation phenotypes by light microscopy. This result shows that ADP1 is relatively likely to non-adaptively evolve this phenotype, possibly due to the large number of genes involved in extracellular polysaccharide biosynthesis or mutational hot- spots affecting those genes. Still, aggregation did evolve at a significantly greater rate in the AE experiment compared to the MA experiment ($P = 0.003$ by chi-squared test comparing Poisson rate models), suggesting that mutations leading to this phenotype in the AE experiment are beneficial to fitness.

Prophage region deletion hotspot

The sequence of the new junction created by the 49 kb LPD, which was present in all seven of the sequenced clones from the AE experiment, suggests that excision of this genomic region was facilitated by recombination between nearly identical flanking 51 bp repeats (**Figure 2.8**). To determine if an elevated rate for this mutation could wholly explain this extreme genetic parallelism, we used PCR to assay for this mutation in the 18 ADP1 MA lines that were evolved with reduced selective pressure for ~7,500 generations. The LPD was not found in any of the MA clones. Thus, we conclude that

this large deletion is a beneficial mutation under the AE conditions because it occurs at a significantly greater rate than that in the MA experiment ($P = 9 \times 10^{-11}$ by chi-squared test comparing Poisson rate models).

The LPD results in the loss of three genes (*ACIAD2139*, *ACIAD2143*, and *ACIAD2190*) reported to be essential for growth in minimal succinate (MS) medium in a previous study of the ADP1 single-gene deletion strain collection (101). However, we found that all sequenced AE clones with the LPD still were able to grow to high density in liquid MS medium and form colonies on MS agar within 24 h. A number of nonessential genes that might affect fitness in the environment of the AE experiment also reside in the LPD region, including other phage-related genes, a putative type II toxin-antitoxin system (*ACIAD2181-ACIAD2182*), a gene encoding a predicted threonine efflux protein (*ACIAD2144*), and a gene encoding a predicted cell wall-associated hydrolase (*ACIAD2149*).

DISCUSSION

In this study, we examined the stability of the undomesticated *A. baylyi* ADP1 genome as it adapted to growth in a laboratory environment over 1,000 generations. We concentrated on characterizing three types of genetic changes that commonly evolved in this strain that affected major traits: loss of its high natural transformability, evolution of cellular aggregation, and deletion of a cryptic prophage region. We did not attempt to directly measure the effects of these mutations on the competitive fitness of ADP1, because the transformability and aggregation phenotypes are expected to lead to complex frequency-dependent effects, such that their benefits depend on the composition of phenotypes in the population in which they evolved. Instead, we inferred that these

mutations or phenotypes were beneficial for fitness in the laboratory environment because we never or rarely observed their evolution under nonselective genetic drift conditions during a mutation accumulation experiment that lasted ~7,500 generations.

IS*1236*, the only native transposable element in the *A. baylyi* ADP1 genome, was the primary cause of mutations affecting transformability and aggregation. Based on sequence homology, IS*1236* has been assigned to the IS3 transposon family (120). Therefore, it is assumed to undergo conservative replication by a copy-and-paste mechanism (121). IS*1236* is present in six copies in the ancestral strain (106), two of which flank the composite transposon Tn*5613* (116).

IS*1236* has been found to play an important role in a number of other studies of ADP1 metabolism and evolution. For example, Tn*5613* was the primary cause of loss-of-function mutations in the *vanA* and *vanB* genes, which are required for the metabolism of ferulate and vanillate (116). Similarly, studies of p-hydroxybenzoate metabolism showed that IS*1236* accounted for ~85% of the mutations inactivating one gene in the pathway, *pobR* (120). IS*1236* also has been implicated in facilitating gene amplifications, including 86% of the amplification events observed in a region located close to Tn*5613* (122). Sequencing of a chemostat population evolved under nitrogen-limiting conditions for ~2,800 generations found five IS*1236* insertions and several large deletions and genomic rearrangements among 22 total mutations profiled (123), none of which affected the genes commonly mutated in our AE experiment. The abundance of both new copies of IS*1236* and deletions and other chromosomal rearrangements mediated by this mobile genetic element demonstrates its central role in changing strain function and genome structure.

We observed unexpected differences in the rates of IS*1236*-mediated mutations among our sequenced clones. At least two different types of effects could have

contributed to this variation. First, early insertions of new copies of an IS element might have led to an increase in transposase expression or IS DNA template availability that snowballed into an increased rate of further IS*1236* transposition events within a lineage. Second, ancillary mutations at another genetic loci could have resulted in increased IS activity due to trans-regulatory effects. The latter explanation appears to be more likely, as the five sequenced clones that have the most IS activity all experienced mutations affecting the ClpA ATPase subunit of the Clp protease. ClpX, another Clp-associated ATPase, has been found to regulate Mu transposition by degrading its Rep repressor protein in *E. coli* cells (124). Similarly, it is possible that the clpA mutations we observed increase Clp-dependent degradation of the IS*3*-type OrfA repressor protein (125) encoded within IS*1236* and thereby hyperactivate its transposition.

If the *clpA* mutations result in increased IS*1236* transposition, it is theoretically possible that parallel evolution of these mutations was the result of second-order selection for increased evolvability associated with this hypermutator effect (105). However, this kind of second-order benefit of leading to an increase in potential beneficial mutations seems unlikely to have been sufficient for the *clpA* mutations to reach such prominence on the time scale of our experiment. Most probably, the *clpA* mutations were directly beneficial for growth in LB, and their effect on transposition, if any, was incidental. ClpA activity will affect the expression of its multiple protein substrates, and adaptive mutations in these types of regulatory hub proteins are common early in laboratory evolution experiments (126, 127). For example, a mutation in *hslU*, a similar ATPase component of the separate HslVU protease, occurs in the Lenski *E. coli* evolution experiment (128).

We found that all of our adapting populations lost transformability over the course of the evolution experiment. A similar result was observed in a previous experiment by

Bacher et al., where transformation frequencies were reduced by 80 to 85% after propagating ADP1 in LB for ~730 generations (98). We observed the progressive displacement of competent isolates by noncompetent mutants within one experimental population that we profiled over several hundred generations. Thus, the greater loss of transformability in our experiment (ranging from 85 to 99% in different populations) likely can be attributed simply to its longer duration, giving more time for newly evolved ADP1 variants carrying these apparently beneficial mutations to outcompete their ancestors and competitors. It is also possible that a molecular drive mechanism contributes to the spread of these loss-of-competence alleles: DNA from cells with a nonfunctional competence gene has a greater chance of being taken up and integrated into the genomes of their highly competent competitors than vice versa (98).

Investigating the genetic basis of reduced transformation in our focal evolved clones showed that IS*1236* insertions that interrupted the known competence genes *pilB* and *comC* were responsible for this phenotype. The much lower incidence of loss-of-competence mutations in our mutation accumulation lines provides evidence that this phenotype is beneficial for ADP1 during competitive growth in nutrient broth. Previously, it has been demonstrated that adding ADP1 genomic DNA to a culture inhibits the growth of wild-type ADP1 more than it inhibits the growth of a noncompetent ADP1 mutant (98). As *A. baylyi* ADP1 also is known to release significant quantities of genomic DNA via lysis of 1 to 5% of the cells in a culture during normal laboratory growth in LB (129), it seems likely that DNA released from lysed cells within our evolving ADP1 populations creates conditions where noncompetent mutants have a fitness advantage.

In addition to the loss of competence, we found that IS*1236*-mediated mutations often were responsible for increased aggregation in many of our evolved clones, because

they disrupted genes involved in extracellular polysaccharide production. This is a particularly interesting observation considering the history of the ADP1 strain. *Acinetobacter baylyi* ADP1, previously designated *Acinetobacter calcoaceticus* BD413 (26), was derived from strain BD4, whose large polysaccharide capsule led to levels of cellular aggregation that were problematic in laboratory cultures (28). BD413 was generated as a nonclumping mutant of BD4, a feature attributed to its "mini encapsulation" phenotype. In the same study, the complete loss of the extracellular polysaccharide capsule was found to cause cellular aggregation (28).

We suspect that the disruption of *per* and *pgi* activity in the evolved mutants results in a similarly severe deficiency in extracellular polysaccharide production. This hypothesis is supported by a concomitant reduction in the activity of a bioemulsifier, which consists of a protein component complexed to these polysaccharides, in the aggregating mutants. In wild-type ADP1, the presence of this emulsifier and/or residual extracellular polysaccharides on the cell surface may prevent clumping. The *per* and *pgi* mutations that evolved separately in one sequenced clone each yielded apparently identical phenotypes with respect to aggregation and bioemulsifier production on their own. Because these mutations were always found together in evolved strains, it seems likely that there was some benefit to having both mutations over just one, perhaps due to other effects they have on cellular physiology or metabolism.

The ~49 kb large prophage deletion (LPD) found in all of our AE experiment populations removes the larger of the two putative prophages that exist in the genome of the ancestral strain (106). Its size was previously estimated as ~52 kb, and the majority of its genes are unique to *A. baylyi*. The homologous terminal repeats that apparently are involved in a recombination event leading to the LPD likely represent the original integration site of the phage, as the downstream repeat copy partially overlaps a tRNA

gene, a common integration site for phage (130). The P22 phage attachment site in Salmonella, for instance, is a 46 bp sequence within the threonine tRNA gene. High frequency loss of the phage region may have resulted from residual activity of a cryptic prophage integrase/recombinase. A putative copy of such a gene is located within the deleted region adjacent to the downstream terminal repeat. Alternatively, the LPD also could have occurred through a host-mediated recombination mechanism.

The rate at which the large prophage deletion occurs during ADP1 growth, as estimated from the MA experiment, is too low to explain its ubiquity in the sequenced AE clones. Thus, it is likely that the LPD was advantageous to fitness under these experimental conditions. Studies of ribose operon deletions in the Lenski long-term *E. coli* evolution experiment exemplify how a locally elevated rate of mutations coupled with even just a slight fitness benefit for those mutations can rapidly lead to extreme genetic parallelism in this manner (131). The LPD contains three putative phage-related genes identified as essential for growth in minimal medium, two of which are predicted DNA-binding regulatory proteins. As we observed that all of the sequenced clones with the LPD still were capable of robust growth in minimal medium, it is possible that deleting these genes singly results in a lethal reactivation of toxic gene products within this latent prophage or other deleterious genetic interactions that do not apply when the ~60 genes in the LPD are excised simultaneously from the genome. It is also possible that a fitness benefit could be conferred by the loss of nonessential genes within the LPD.

Overall, our study highlights several challenges that may be encountered when engineering or evolving the genome of any undomesticated bacterial strain in the laboratory. Specifically regarding ADP1 genome engineering, we note that the LPD also removes the genomic location targeted by a toolkit of integration vectors (104). In this case, the choice of the large prophage region as a neutral site for inserting heterologous

genes into the ADP1 genome should be reconsidered in light of its inherent genetic instability, as this deletion was the most prevalent mutation observed in ADP1 propagated in LB. This discovery, combined with the overall high rate of transposon-mediated mutations, the rapid loss of transformability, and the atavistic reappearance of cellular aggregation, demonstrates that the native features of the ADP1 genome currently present significant challenges to using this strain as a platform for genome engineering.

However, ADP1 has many potential advantages as a versatile host microorganism for synthetic biology (23, 24, 94), especially considering recent studies describing multiplex genome editing in naturally transformable organisms (87). The genetic stability of ADP1 could be improved by domesticating its genome by deleting IS elements, cryptic prophages, potential toxin-antitoxin cassettes, and other uncharacterized genes that have accumulated during its prior existence in the wild outside the laboratory, as has been done in the "clean-genome" *E. coli* MDS42 strain (92). This scale of refactoring could be accomplished relatively easily in ADP1 with existing genetic tools (24), particularly given that there are only six copies of IS*1236* in the genome. Metabolic pathways and complex biological devices assembled in a clean-genome ADP1 strain would be expected to exhibit greater genetic reliability on evolutionary timescales (132), and such a host could enable more facile genome engineering for synthetic and evolutionary biology studies involving widespread gene acquisition and loss (91).

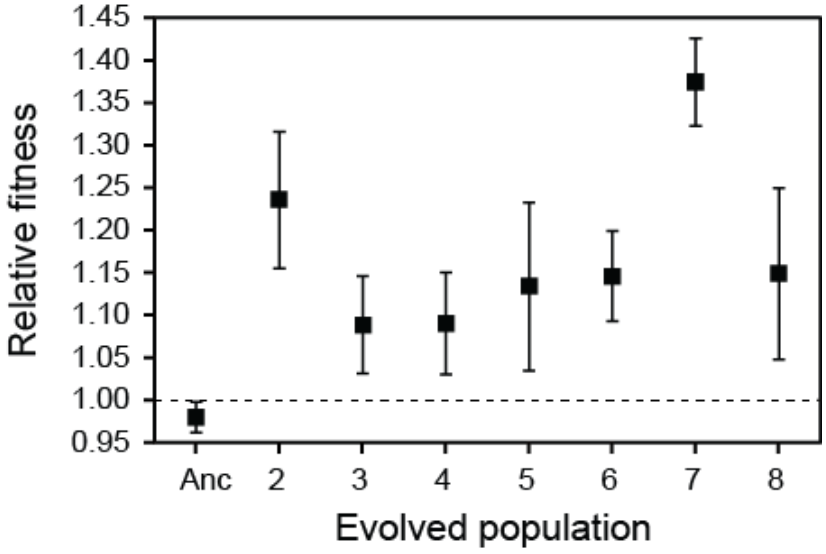

Figure 2.1. *A. baylyi* ADP1 evolved increased fitness during the 1,000-generation adaptive evolution experiment.

Co-culture competition assays were conducted between a GFP-tagged variant of the ancestor and the ancestral strain or population samples. Error bars are 95% confidence limits estimated from at least eight replicate assays.

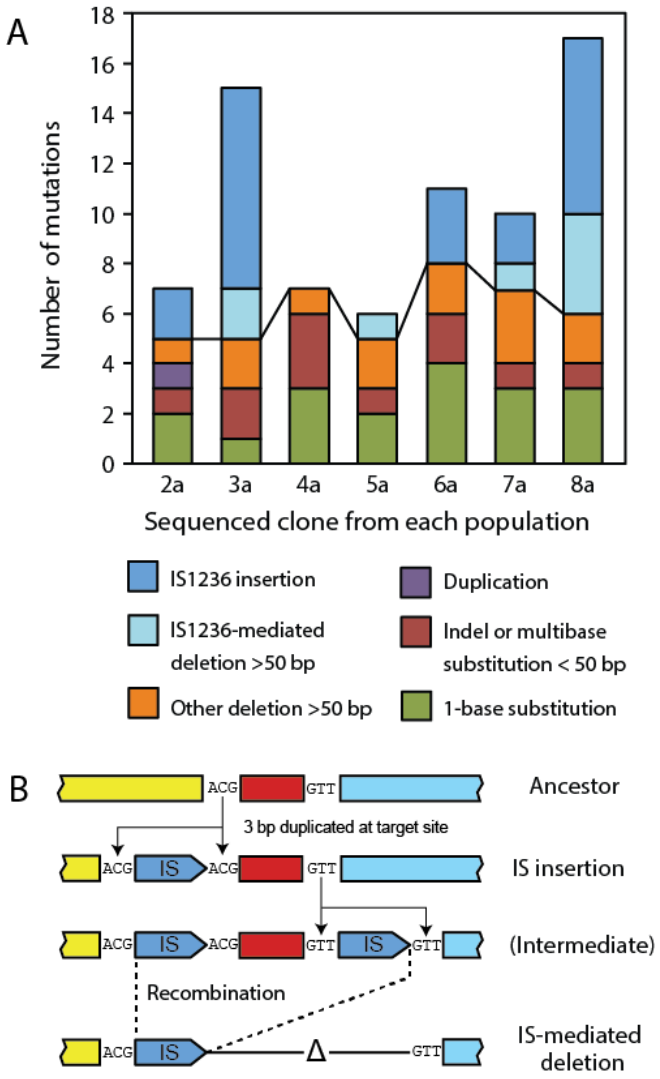

Figure 2.2. Genome evolution during the 1,000 generation *A. baylyi* ADP1 adaptive evolution experiment.

(A) Types of mutations observed in the sequenced clones from each of the seven independently evolved populations. Lines connecting the columns divide mutations caused by new IS*1236* activity from other categories of mutations. Deletions caused by homologous recombination between the two ancestral copies of IS*1236* that flank the transposon Tn*5613* were not counted as IS-mediated. Full details for each mutation are provided in **Table 2.2**. (B) Schematic of an IS*1236* insertion and an IS*1236*-mediated deletion. A copy-and-paste transposition event results in the duplication of several base pairs (usually three) at the target site flanking the new IS*1236* copy. A second, close-by IS*1236* insertion can lead to a homologous recombination event between the two IS copies that excises the intervening genomic region. One IS*1236* element copy and an adjacent deletion remain after this type of IS-mediated deletion.

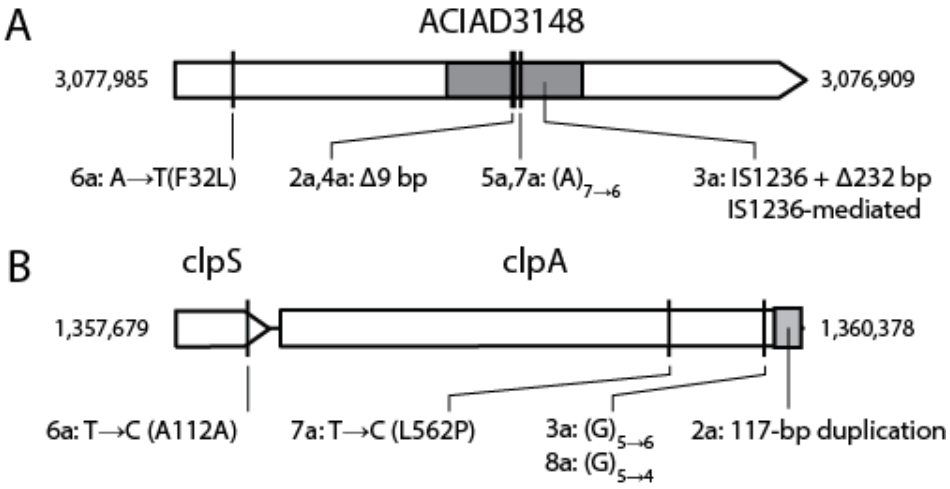

Figure 2.3. *ACIAD3148* and *clpA* genes are mutated in isolates from many populations of the adaptive evolution experiment.

(A) Mutations affecting *ACIAD3148* likely evolved in parallel in six populations because it is beneficial to disrupt this hypothetical protein of unknown function that is found only in the genus Acinetobacter. The mutations observed include a IS*1236* insertion followed by a IS*1236*-mediated deletion, single-base deletions in a run of seven adenine bases, and in-frame deletions of nine base pairs that could be mediated by a five-base repeat. The resulting amino acid change is shown in parentheses for single-base substitutions. (B) Most mutations affecting the ClpA ATPase and specificity component of the Clp protease disrupt only its C-terminal sequence, but there is also a nonsynonymous base substitution affecting amino acid 562 and a mutation that may affect a putative transcriptional promoter for *clpA* located inside the *clpS* gene.

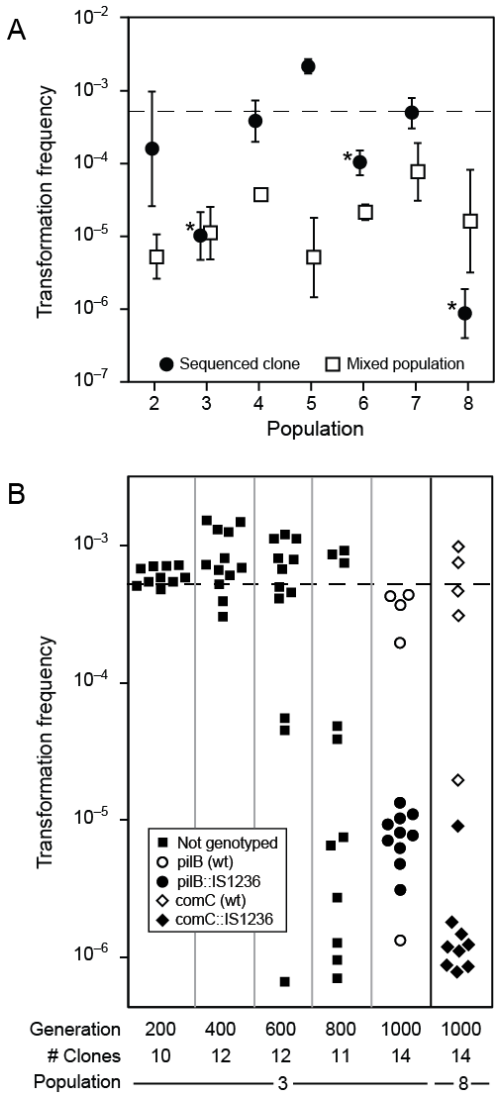

Figure 2.4. IS*1236* drives the evolutionary loss of ADP1 natural transformation.

Transformation frequencies for all evolved populations from the adaptive evolution experiment (unfilled squares) were reduced compared to the wild-type ADP1 ancestor (dashed line). However, for the sequenced clones from each population (filled circles) transformability was only significantly reduced for certain isolates (*$p < 0.05$, one-tailed Welch's *t*-test). Error bars are 95% confidence limits for triplicate assays. (B) In population 3, loss of transformation ability is mediated by the expansion of a non-competent subpopulation over time as determined by testing individual clones picked from the mixed population. For populations 3 and 8, several 1,000-generation clones were genotyped for the relevant IS*1236* insertions in competence genes found in the corresponding sequenced clone. In each case, these IS*1236* insertions appear to be responsible for most of the loss of competence in the population although other, rarer mutations that reduce competence also appear to be present in each case. The measurements shown are for triplicate assays. Error bars are omitted for clarity.

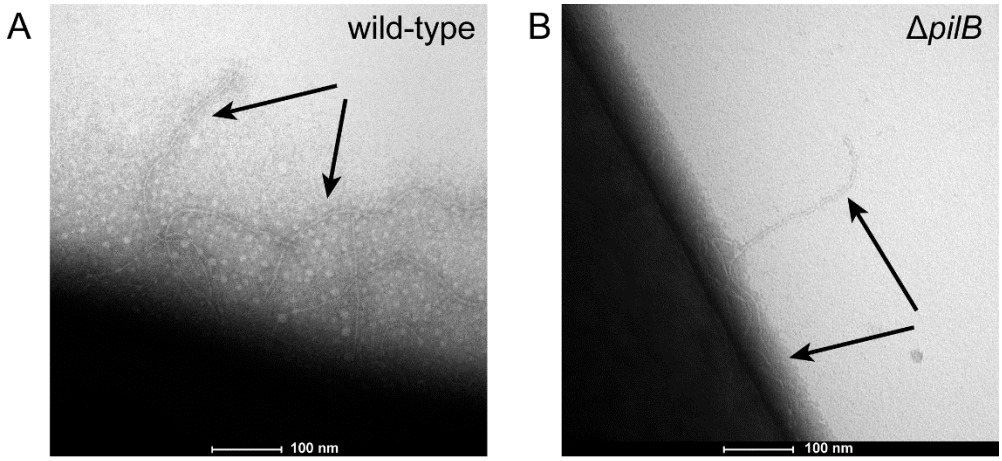

Figure 2.5. Transmission electron micrographs of *A. baylyi* ADP1

Visualized are both (A) wild-type *A. bayli* ADP1 and (B) the Δ*pilB* mutant. In both cases, negative staining with uranyl acetate was used to visualize pili extending from the surfaces of cells (indicated by arrows).

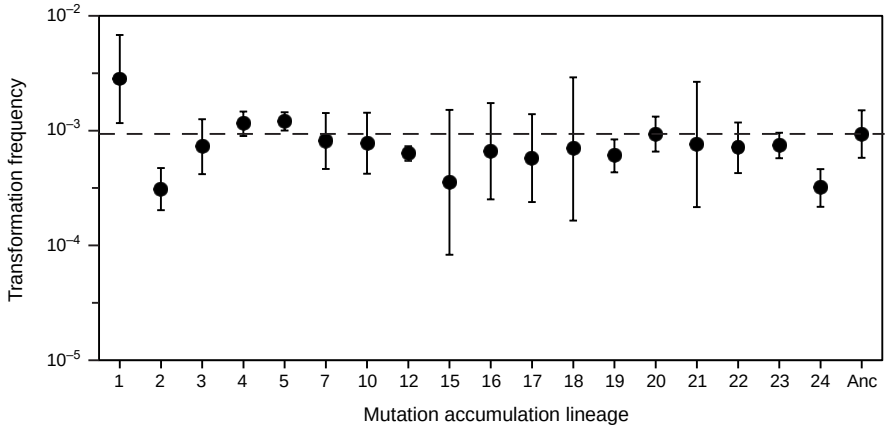

Figure 2.6. Transformation frequencies were measured for 7,500-generation endpoint clones from 18 lineages of the mutation accumulation experiment.

Error bars are 95% confidence limits for triplicate assays. The dashed line corresponds to the transformation frequency of the ancestral ADP1 strain. None of the tested clones had a transformation frequency that was significantly reduced below half that of the ancestor ($p > 0.05$, Bonferroni-corrected one-tailed Welch's *t*-test).

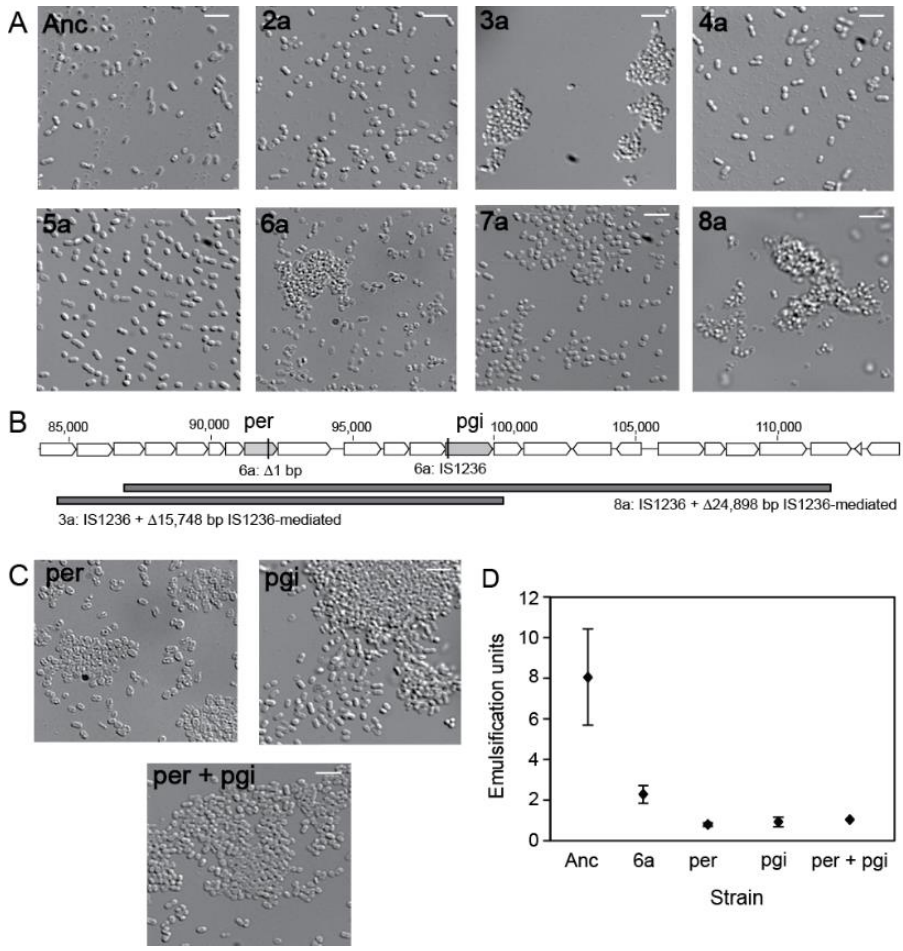

Figure 2.7. IS*1236*-mediated loss of genes related to extracellular polysaccharide (EPS) production leads to increased cellular aggregation and reduced bioemulsification activity in evolved *A. baylyi* ADP1 strains.

(A) Sequenced clones have two distinct phenotypes visible by light microscopy: clones with the ancestral phenotype (Anc) occur in clusters of no more than two cells (2a, 4a, 5a, 7a) and clones with a high-aggregation phenotype are typically found in clusters of tens to hundreds of cells (3a, 6a, 8a). Scale bars represent 5 μM. (B) The three sequenced clones with the high-aggregation phenotype have loss-of-function mutations in the *per* and *pgi* genes related to EPS production. In two cases, these genes are located within large IS*1236*-mediated deletions. In clone 6a, each gene is inactivated by a separate mutation. (C) Reconstruction of the *per* and *pgi* mutations from clone 6a in the ADP1 ancestor shows that either mutation alone confers the high aggregation phenotype and that the double mutant does not exhibit greater aggregation than either single mutant. Scale bars represent 5 μM. (D) Loss of either *per* or *pgi* function in strains with the mutations in the 6a clone similarly leads to ADP1 cultures with reduced activity for emulsifying a water-hydrocarbon mixture. Error bars are S.E.M. of triplicate assays.

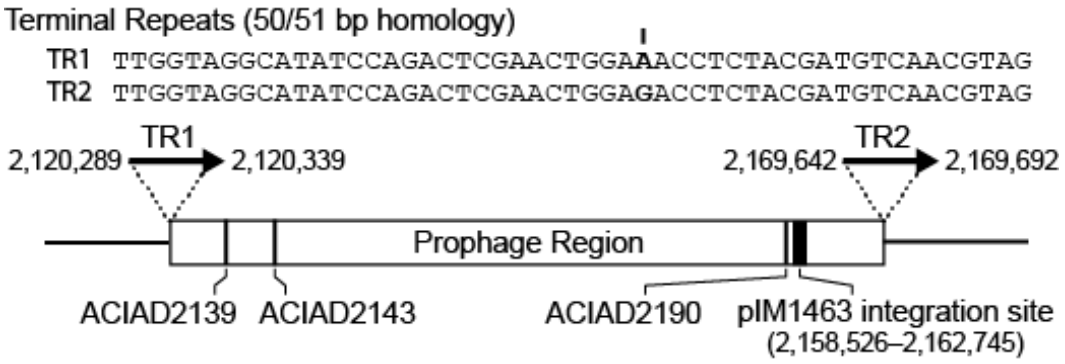

Figure 2.8. Terminal repeats flanking the large prophage region of *A. baylyi* ADP1 appear to mediate its rapid loss during the adaptive evolution experiment.

The one base pair difference between the two repeats is bolded and marked for emphasis. Deletion of this unstable region in all seven sequenced clones results in the loss of three prophage-related genes previously defined as essential for growth in minimal media (101) — *ACIAD2139*, *ACIAD2143*, and *ACIAD2190* — as well as a common site for genomic modification targeted by integration vectors such as pIM1463 (133).

Position	Mutation	Annotation	Gene	Description
178,468	Δ1 bp	intergenic (+100/- 74)	*znuA* → / → *atpI*	high affinity Zn ABC transporter periplasmic substrate- binding protein/ATP synthase protein I
1,179,108	C→T	D388N (GAC→AAC)	*phrB* ←	deoxyribodipyrimidine photolyase (photoreactivation), FAD- binding
1,278,410	A→C	T195P (ACC→CCC)	*ACIAD1277* →	hypothetical protein
†1,803,858	C→T	I104I (ATC→ATT)	*ACIAD1796* →	hypothetical protein
†1,803,861	T→C	A105A (GCT→GCC)	*ACIAD1796* →	hypothetical protein
†1,803,900	C→T	V118V (GTC→GTT)	*ACIAD1796* →	hypothetical protein
†1,803,903	T→C	Y119Y (TAT→TAC)	*ACIAD1796* →	hypothetical protein
†1,803,915	G→A	K123K (AAG→AAA)	*ACIAD1796* →	hypothetical protein
†1,803,921	A→G	A125A (GCA→GCG)	*ACIAD1796* →	hypothetical protein
†1,803,930	T→C	D128D (GAT→GAC)	*ACIAD1796* →	hypothetical protein
†1,803,933	A→G	L129L (TTA→TTG)	*ACIAD1796* →	hypothetical protein
†1,803,939	C→T	G131G (GGC→GGT)	*ACIAD1796* →	hypothetical protein
2,420,279	+C	intergenic (+56/+34)	*ACIAD2457* → / ← *glnA*	hypothetical protein/ glutamine synthetase
2,542,652	G→A	I19I (ATC→ATT)	*lepA* ←	GTP- binding protein LepA
2,765,010	A→C	S353R (AGT→CGT)	*ACIAD2827* →	periplasmic binding protein of transport/transglycosylase
¶2,803,338	1,890 bp × 2	duplication	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,803,531	G→A	G930G (GGC→GGT)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,910	A→G	S471P (TCA→CCA)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,911	+324 bp*	coding (1410/5178 nt)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,914	A→G	T469T (ACT→ACC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,917	A→G	D468D (GAT→GAC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,938	A→G	S461S (AGT→AGC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,956	T→C	G455G (GGA→GGG)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,804,983	G→A	I446I (ATC→ATT)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,028	A→G	S431S (AGT→AGC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,037	A→G	N428N (AAT→AAC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,054	T→C	I423V (ATC→GTC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,067	G→A	T418T (ACC→ACT)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,145	G→A	T392T (ACC→ACT)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
‡2,805,196	A→G	I375I (ATT→ATC)	*ACIAD2866* ←	hemagglutinin/hemolysin- related protein
§3,075,879	G→A	noncoding (582/1530 nt)	*ACIADrRNA16S_18* ←	16S rRNA
3,357,842	G→A	intergenic (- 237/- 520)	*ACIAD3437* ← / → *ACIAD3440*	hypothetical protein/hypothetical protein

Table 2.1 Discrepancies between the ancestral *A. baylyi* ADP1 strain and the reference genome.

Position: Coordinates in the Acinetobacter baylyi ADP1 genome sequence (Genbank: NC_005966.1) of the sequence change.
Mutation: Description of the sequence difference.
Annotation: For single-base mutations in protein coding genes, the effect of the sequence change on the amino acid sequence and codon. For intergenic mutations, the distances to the two nearest flanking genes, separated by a slash, with negative values indicating that the mutation is upstream and positive values meaning it is downstream of the respective gene.

Gene: Name or ID of affected gene or genes with arrows indicating the genomic strand. For intergenic mutations, the identifiers for the two nearest flanking genes, separated by a slash.
Description: Functional annotations for affected gene products.
† *ACIAD1796* gene conversion consisting of 9 base changes.
¶ The presence of the *ACIAD2866* duplication in the ancestral strain was verified by PCR across its boundary.
‡ *ACIAD2866* gene conversion consisting of 13 base changes and one 324 base pair insertion. These changes only affect the first copy of the 1,890 bp duplication.
§ 16S gene conversion consisting of a single base change.
* The new sequence inserted after this position is:

GTGGTGTCTACAGTGAAGCTGAGTTCAGTTGATGTCGTGCCAACATTGCCTGCTTGGTCCACAATCGACGCTTTGTAGT
TGTGGGTGCCTTCGCTGACATTGCTCAATTGGTAGCTGAAGCTTGCGCCTGTGACGGTCGCTGTACCCAGTAAGGTGTT
GCCTTCATAGATCTGAACCACATCACCATCTTTAAGACCACTCACTGTACCTTTGAGTAGTGGTGAGGTGTCATTGGTC
GATGTGCCTGTACCGTAGTCACCGGTATTCGGTGCAACGTCATCAGTATAGCCTGTCAGGCTGATGCTGTTATCTGTGG
TTGGCGCT

Table 2.1: continued

Position	Mutation	Clones	Annotation	Gene	Description
84,583	IS*1236* (+) +3 bp	3a	coding (336- 338/1032 nt)	*ACIAD0088* →	NAD-dependent epimerase/dehydratase
84,584	Δ15,748 bp	3a	IS*1236*- mediated	*[ACIAD0088]–[galE]*	*[ACIAD0088], ACIAD0089, ACIAD0090, ACIAD0091, ACIAD0092, ACIAD0093, ACIAD0094, per, ACIAD0096, ACIAD0098, galU, ACIAD0100, pgi, [galE]*
86,945	IS*1236* (–) +3 bp	8a	coding (375- 377/1101 nt)	*ACIAD0090* →	glycosyl transferase
86,947	Δ24,898 bp	8a	IS*1236*- mediated	*[ACIAD0090]–[ACIAD0110]*	*20 genes [ACIAD0090], ACIAD0091, ACIAD0092, ACIAD0093, ACIAD0094, per, ACIAD0096, ACIAD0098, galU, ACIAD0100, pgi, galE, ACIAD0103, manB, ACIAD0105, lldP, lldR, lldD, dld, [ACIAD0110]*
92,042	Δ1 bp	6a	coding (844/1176 nt)	*per* →	perosamine synthetase
98,370	+TT :: IS*1236* (+) +1 bp	6a	coding (66/1674 nt)	*pgi* →	glucose-6-phosphate isomerase
153,626	IS*1236* (+) +3 bp	8a	coding (1015- 1017/2481 nt)	*ACIAD0152* ←	hypothetical protein
282,061	T→G	6a	T218P (ACC→CCC)	*argH* ←	argininosuccinate lyase
322,810	Δ3,510 bp	5a	IS*1236*- mediated	*[ACIAD0323]–[ACIAD0330]*	*[ACIAD0323], ACIAD0324, ACIAD0328, [ACIAD0330]*
322,810	Δ892 bp	8a	IS*1236*- mediated	*ACIAD0323* ←	LysR family transcriptional regulator
356,110	IS*1236* (–) +3 bp	3a	coding (256- 258/1740 nt)	*pilB* ←	type 4 fimbrial biogenesis protein
938,749	IS*1236* (+) +3 bp	8a	coding (682- 684/843 nt)	*ACIAD0947* →	hypothetical protein
942,544	Δ2,315 bp	3a, 6a, 7a, 8a	between IS*1236*	*ACIAD0955–[ACIAD0959]*	*ACIAD0955, ACIAD0956, ACIAD0957, ACIAD0958, [ACIAD0959]*
944,858	Δ2 bp	6a	coding (761- 762/765 nt)	*ACIAD0959* ←	hypothetical protein
1,122,949	C→T	8a	A360A (GCG→GCA)	*ACIAD1136* ←	hypothetical protein
1,358,044	T→C	6a	A122A (GCT→GCC)	*clpS* →	ATP-dependent Clp protease adaptor protein ClpS
1,359,786	T→C	7a	L562P (CTT→CCT)	*clpA* →	ATP-binding protease component
1,360,199	$(G)_{5\rightarrow4}$	8a	coding (2098/2277 nt)	*clpA* →	ATP-binding protease component
1,360,199	$(G)_{5\rightarrow6}$	3a	coding (2098/2277 nt)	*clpA* →	ATP-binding protease component
1,360,244	117 bp x 2	2a	duplication	*clpA* →	ATP-binding protease component
1,445,209	A→C	2a	L165F (TTA→TTC)	*catB* →	muconate cycloisomerase I
1,576,618	T→C	7a	intergenic (- 115/+140)	*srpI* ← / ← *srpH*	major membrane protein I (MMP- I) (35 kDa antigen)/serine acetyltransferase
1,721,316	A→G	5a	T44A (ACC→GCC)	*quiA* →	quinate/shikimate dehydrogenase [pyrroloquinoline- quinone] (NAD(P)-independent quinate dehydrogenase)
1,777,498	IS*1236* (–) +3 bp	6a	intergenic (+206/- 96)	*ACIAD1769* → / → *ACIAD1770*	branched-chain amino acid ABC transporter ATP- binding protein/hypothetical protein
1,778,255	Δ21,347 bp	8a	IS*1236*- mediated	*ACIAD1772–[ACIAD1790]*	*18 genes ACIAD1772, ACIAD1773, otsA, otsB, ACIAD1776, fecI, fecR, ACIAD1780, ACIAD1781, ppk, ACIAD1783, ACIAD1784, ACIADtRNAMet_44, ACIAD1786, csp, ACIAD1788, ACIAD1789, [ACIAD1790]*
1,977,056	Δ1 bp	4a	coding (355/444 nt)	*ACIAD1986* ←	hypothetical protein

Table 2.2 Mutations identified in ADP1 isolates from the adaptive evolution experiment

Position	Mutation	Clones	Annotation	Gene	Description
1,988,361	Δ1,125 bp	5a		*[ACIAD1997]–[dadA]*	*[ACIAD1997], [dadA]*
1,988,376	IS*1236* (–) +3 bp	7a	coding (379- 381/651 nt)	*ACIAD1997* →	hypothetical protein
2,014,513	IS*1236* (–) +3 bp	8a	coding (795- 797/831 nt)	*ACIAD2022* →	AraC family transcriptional regulator
2,099,589	IS*1236* (–) +3 bp	3a	intergenic (- 571/- 37)	*ACIAD2109* ← / → *ACIAD2110*	hypothetical protein/hypothetical protein
2,120,340	Δ49,353 bp	2a, 3a, 4a, 5a, 6a, 7a, 8a	Large Prophage Deletion	*ACIAD2132–[ACIADtRNAVal_46]*	*63 genes ACIAD2132– ACIAD2200, [ACIADtRNAVal_46]*
2,248,854	IS*1236* (–) +3 bp	3a	coding (127- 129/390 nt)	*ACIAD2278* →	4-carboxymuconolactone decarboxylase
2,478,738	A→G	7a	intergenic (- 98/+265)	*ACIAD2521* ← / ← *ACIAD2522*	hypothetical protein/transcriptional regulator
2,478,767	T→G	2a, 3a, 6a, 8a	intergenic (- 127/+236)	*ACIAD2521* ← / ← *ACIAD2522*	hypothetical protein/transcriptional regulator
2,580,176	A→T	4a	I174F (ATC→TTC)	*ACIAD2627* →	RNA polymerase factor sigma-70
2,580,921	IS*1236* (+) +3 bp	2a	coding (303- 305/372 nt)	*ACIAD2629* →	hypothetical protein
2,725,447	Δ1,009 bp	3a*		*[ACIAD2778]*	*[ACIAD2778]*
2,725,835	Δ620 bp	7a	intergenic (- 122/+6)	*ACIAD2778* ← / ← *ACIAD2779*	AraC family transcriptional regulator/hypothetical protein
2,748,474	IS*1236* (–) +3 bp	3a	coding (202- 204/300 nt)	*ACIAD2804* ←	hypothetical protein
2,754,073	Δ995 bp	8a	IS*1236*- mediated	*[ACIAD2817]*	*[ACIAD2817]*
2,754,073	IS*1236* (–) +3 bp	8a	intergenic (- 13/+341)	*ACIAD2815* ← / ← *ACIAD2817*	hypothetical protein/transcriptional regulator
2,755,073	IS*1236* (+) +3 bp	3a	coding (103- 105/762 nt)	*ACIAD2817* ←	transcriptional regulator
2,812,774	2 bp→TT	3a, 4a	coding (659- 660/891 nt)	*sucD* ←	succinyl-CoA synthetase subunit alpha
3,000,184	IS*1236* (+) +3 bp	3a	coding (1719- 1721/2799 nt)	*barA* →	GacS-like sensor kinase protein
3,000,834	IS*1236* (–) +3 bp	8a	coding (2369- 2371/2799 nt)	*barA* →	GacS-like sensor kinase protein
3,077,395	$(A)_{7\rightarrow6}$	5a, 7a	coding (591/1077 nt)	*ACIAD3148* ←	hypothetical protein
3,077,406	Δ9 bp	2a, 4a	coding (572- 580/1077 nt)	*ACIAD3148* ←	hypothetical protein
3,077,521	Δ232 bp	3a	IS*1236*- mediated	*ACIAD3148* ←	hypothetical protein
3,077,521	IS*1236* (–) +3 bp	3a	coding (463- 465/1077 nt)	*ACIAD3148* ←	hypothetical protein
3,077,890	A→T	6a	F32L (TTT→TTA)	*ACIAD3148* ←	hypothetical protein
3,223,721	IS*1236* (+) +3 bp	8a	coding (970- 972/4353 nt)	*comC* ←	competence factor involved in DNA binding and uptake
3,319,140	A→C	4a	L159W (TTG→TGG)	*ACIAD3402* ←	cold shock protein
3,385,496	A→G	5a	F348L (TTT→CTT)	*ACIAD3460* ←	glycine betaine transporter
3,389,989	IS*1236* (–) +3 bp	2a	coding (55- 57/333 nt)	*ACIAD3464* ←	hypothetical protein
3,392,671	IS*1236* (+) +2 bp	6a	coding (2477- 2478/3531 nt)	*ACIAD3465* →	two-component sensor
3,394,195	Δ2,988 bp	7a	IS*1236*- mediated	*[ACIAD3468]–[ACIAD3471]*	*[ACIAD3468], ACIAD3469, msuE, [ACIAD3471]*
3,394,195	IS*1236* (+) +3 bp	7a	coding (96- 98/768 nt)	*ACIAD3468* →	hypothetical protein
3,395,084	A→C	8a	N169K (AAT→AAG)	*ACIAD3469* ←	two-component response regulator
3,422,362	C→T	4a	M119I (ATG→ATA)	*prfB* ←	peptide chain release factor 2

Table 2.2 Mutations identified in ADP1 isolates from the adaptive evolution experiment (continued)

Column descriptions are as in **Table 2.1**. * The frequency of this deletion in the sample was ~70% based on counts of reads mapping across the junction and reduced sequencing coverage of this region, indicating that it may have occurred during colony outgrowth and that the sample may not have been purely clonal. This deletion was not counted as a consensus mutation in **Figure 2.1**.

Chapter 3: Emergence of a competence reducing filamentous phage from the genome of *Acinetobacter baylyi* ADP1

INTRODUCTION

Prophages, genomically integrated copies of viral DNA, are common in bacteria. They can comprise significant portions of bacterial genomes (134) and critically impact host cells (135). Some prophages are derived from recent infections with temperate phage and are still capable of replicating and re-infecting cell populations if induced stochastically or by environmental triggers. However, many prophages are believed to be cryptic, having accumulated mutations that disrupt their ability to produce infective phage particles. Both intact and cryptic prophages can confer benefits to their hosts, such as by providing proteins that aid in stress responses (136), by selectively killing lysogen-free competitors (137), or by promoting horizontal gene transfer (138). Prophages can also carry virulence factors that are necessary for bacteria to adopt pathogenic lifestyles, as is the case for the CTXϕ virus of *Vibrio cholera* (139).

Naturally competent bacteria are capable of transforming DNA without any artificial chemical or electrical treatments (140). DNA uptake is most often accomplished through the use of a type IV pilus structure that binds and translocates extracellular DNA into the cell (141). The ease of transforming designed DNA sequences into naturally competent species makes them attractive targets for genome-scale engineering. For instance, the naturally competent Gram-negative bacterium *Acinetobacter baylyi* ADP1 has been proposed as platform for synthetic biology due to its metabolic versatility and high natural competence (24, 25). Unlike other bacterial species that only become competent under specific environmental conditions or in which only a fraction of cells in a population become competent—such as *Vibrio cholera* (55) or

Bacillus subtilis (142)—all or most ADP1 cells are constitutively competent during exponential growth (143).

We and others have found that ADP1 evolves reduced transformability when it is propagated in the laboratory (98, 144). This result was thought to occur due to a strong selective pressure under these conditions, either a direct cost for producing the DNA uptake apparatus or a deleterious effect from active DNA uptake. However, while studying the evolutionary dynamics in these populations in more detail, we discovered that a filamentous prophage in the ADP1 genome had become reactivated during these experiments. This phage inhibits the growth and transformability of ADP1, but it has no effect on strains deleted for genes required for competence. Therefore, we have named it the Competence Reducing Acinetobacter Phage (CRAϕ). The timing with which CRAϕ emerged in the evolution experiment suggests that loss of ADP1 competence may have been selected because it protects against phage infection.

CRAϕ resembles the CTXϕ bacteriophage from *Vibrio cholerae* in its genetic organization and virion morphology (139). Filamentous phages are known to infect bacteria through type IV pili: for example, the toxin co-regulated pilus of *V. cholerae* in the case of CTXϕ (139) or the conjugative F-pilus of *Escherichia coli* in the case of phages f1 and M13 (145). To the best of our knowledge, CRAϕ is the first phage known to specifically target a type IV pilus that functions in environmental DNA uptake. Prophage sequences with homology to CRAϕ are present in the genomes of other Acinetobacter species, including several *Acinetobacter baumannii* strains. Like CRAϕ, they encode homologs of the *V. cholerae* CTXϕ virulence factors suggesting that these prophage may contribute to *A. baumanii* pathogenicity. We further hypothesize that the rarity of competent bacteria in nature may reflect a balance between the potential benefits of DNA uptake (146, 147) and the risk of infection by bacteriophages like CRAϕ.

MATERIALS AND METHODS

Strain and growth conditions

Strains in this study are derived from an *A. baylyi* ADP1 evolution experiment (144). Growth conditions were incubation at 30°C with orbital shaking at 140 r.p.m. over a 1-inch diameter in LB (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl). Population and clone samples to be studied were revived by inoculating 2 μL of a −80°C glycerol stock (20% v/v glycerol) into 5 mL of LB in a sterile test tube, allowing overnight growth. Then, they were pre-conditioned by transferring 10 μL of this culture into 10 mL of LB in a sterile 50 mL Erlenmeyer flask and growth for 24±1 h. ADP1 clones from evolved populations were isolated by plating portions of pre-conditioned populations diluted in sterile saline on LB-agar (1.6% w/v) plates, growing the plates overnight at 30ºC, and then picking single colonies. ADP1 knockout strains were obtained courtesy of the collection maintained at Genoscope (Évry, France) (101), except the *pilB* knockout was constructed through chromosomal replacement of this gene with a *tdk*-kan^R cassette using methods described elsewhere (144).

Phage isolation

Phage genomic DNA was isolated from standard overnight cultures of clone P5-C and ancestor ADP1 by following an M13 DNA isolation protocol (148). Briefly, 1 mL of culture was pelleted at 15,000 × *g* for 5 min, after which the supernatant was transferred to a new tube where 200 μL of 20% PEG in 2.5 M NaCl was added. After gentle mixing followed by incubation at room temperature for 15 min, phage particles were precipitated by centrifuging the mixture at 15,000 × *g* for 5 min at 4°C. The supernatant was then

removed, leaving a small and visible pellet in the case of the P5-C clone. No pellet was visible for ancestral ADP1. DNA was purified from this pellet by phenol/chloroform extraction followed by ethanol precipitation.

Transformation assay

Measurements of transformation frequency were conducted as previously described (144). Briefly, we combined 1 mL fresh LB, 100 ng AB-KAN genomic DNA, and 70 µL pre-conditioned culture in a sterile test tube, incubated the resulting culture for 18-24 h under standard growth conditions, and then plated dilutions in sterile saline onto selective (Kan) and non-selective (no antibiotic) agar plates. The ratio of CFUs on selective plates to CFUs on non-selective plates after overnight incubation at 30°C was used to determine the frequency of transformation of the kanamycin resistance marker into the ADP1 genome. All transformation assays were carried out in triplicate. Average transformation frequencies and 95% confidence intervals were calculated from log-transformed values.

The ability of a culture's supernatant to inhibit transformation was assayed by combining 500 µL of sterilized spent media using a 0.22 µm filter from an overnight grown culture with 500 µL of fresh LB, 100 ng AB-KAN genomic DNA, and 70 µL preconditioned ancestor ADP1. As before, transformations were incubated for 18-24 hours before plating. Sterile filtered spent supernatant from ancestral ADP1 strain was used as a control. Spent media was collected from cultures grown for 18-24 hours under standard conditions which were then pelleted via centrifugation to collect supernatant. Where appropriate, cell-free filtered supernatant was boiled by immersing a 1.7 mL

Eppendorf tube in a beaker of boiling water for 15 minutes, then cooled to room temperature before use.

Growth curves

To test whether an overnight culture's supernatant inhibited ADP1 growth, 100 μL of a pre-conditioned ADP1 culture was used to inoculate 9 mL of LB mixed with 1 mL of filter sterilized spent media in a 50 mL flask. Each culture was grown under standard conditions for 6 h, removing 1 mL every hour to measure absorbance at 600 nm (OD600). All conditions were tested in triplicate. Boiled supernatant was prepared as described earlier.

Genome sequencing

Genome sequencing of *A. baylyi* ADP1 clones was carried out as previously described (144). Briefly, genomic DNA was isolated from strains of interest and sheared to an average size of ~550 bp using a Covaris S2. Libraries were prepared using a NEBNext prep kit and sequenced on an Illumina HiSeq 2000 instrument at the University of Texas at Austin Genome Sequencing and Analysis Facility (GSAF). Mutations were identified by comparing DNA sequencing reads to the ADP1 genome (GenBank:NC_005966.1) (106) using the *breseq* computational pipeline (107, 149). The same methodology was used to sequence DNA isolated from phage particles isolated from clone P5-C. FASTQ files of DNA sequencing reads have been deposited in the NCBI Sequence Read Archive (SRP074541).

Phage protein analysis

For electron microscopy and mass spectrometry analyses, phage were PEG precipitated from a 1 L culture, as described above, and then further purified to remove cell debris. PEG precipitated pellets were resuspended in 10 mM Tris•HCl (pH 7.6) with 10 mM $MgCl_2$. The phage were then purified by a CsCl density gradient as in Casjens et al. (150). An aliquot of the purified sample was precipitated with TCA, run on a 15% SDS-PAGE gel, and Coomassie stained. The Research Technology Support Facility (RTSF) at Michigan State University performed mass spectrometry analysis on bands extracted from the gel.

Electron microscopy of negatively stained CRAϕ

A 3.5 μL aliquot of purified CRAϕ was applied to a continuous carbon support film (Ted Pella) that had been plasma cleaned in a Fishione model 1020 plasma cleaner. The sample was briefly washed with distilled water, and then stained with 1% aqueous uranyl formate. Micrographs were recorded on a DE-20 camera (Direct Electron, LP, San Diego, CA) in a JEOL 2200FS microscope at a nominal magnification of 30,000 (1.72 Å per pixel) at an acceleration voltage of 200 keV.

Detection of circularized phage episome and mutations in CRAϕ

Briefly, phage episomes were detected by performing PCRs across a junction that is unique to the circularized phage genome. PCRs across the junctions of the genomically located prophage were conducted in parallel to serve as a positive control. PCR reactions were conducted with Taq polymerase in ThermoPol buffer (New England Biolabs) under standard conditions with purified genomic DNA as template. Mutations in CRAϕ isolates

were detected by Sanger sequencing PCR products amplified from genomic DNA purified from evolved ADP1 cells. Sanger traces were aligned against the ADP1 genome sequence using Geneious (version 6.1.6) software (151).

Identification of CRAϕ-like prophage in other Acinetobacter genomes

A microbial TBLASTN of CRAΦ restricted to Acinetobacter genus members with full genomes, homologs are marked as present if they have an E-value lower than 1E-5. No significant homologies were found for *psh*, *cep*, *orfU*, *ace*, *ACIAD1851*, or *ACIAD1850*. Phage genetic structure calls indicate whether the genes whether proximal arrangements of genes were consistent with a CRAϕ-like phage genome. Strains with partial phage structures (i.e. a section of CRAϕ but not an identifiable complete phage structure) were excluded from the results.

RESULTS

Evolved ADP1 secrete a factor that inhibits transformation and growth

We identified CRAϕ while investigating the results of a long-term laboratory evolution experiment with *A. baylyi* ADP1 (144). We noticed a large discrepancy between the transformability of a sample of one population (designated P5), which had been archived after 1000 generations of evolution, and clones isolated from this population. The whole-population transformation frequency was ~100-fold lower than the average transformation frequency of fifteen randomly selected clones when they were tested individually (**Figure 3.1A**). We observed a similar result for a second population (designated P3). Here, the 1000-generation whole-population sample was ~10-fold less

transformable than expected from taking the average of fifteen clone measurements. These observations might mean that the selected clones were not representative of the true mixtures of genotypes in these populations or that rare subpopulations in each mixed sample were reducing the transformability of other cells in the same culture.

Pursuing the second possibility, we hypothesized that a transformation-reducing subpopulation might secrete a diffusible factor that affected DNA uptake by other cells in the same culture. Therefore, we tested whether sterile-filtered supernatant from each of the 15 clones isolated from P5 was able to inhibit transformation of the ancestral, wild-type ADP1 strain. We found that supernatant from one of the 15 clones (designated P5-C) reduced transformation by ~10-fold (**Figure 3.1B**). The addition of P5-C supernatant also markedly reduced the growth rate of the ancestral strain of ADP1 (**Figure 3.1C**). Boiling the P5-C supernatant completely eliminated its ability to inhibit transformation and growth (**Figure 3.1B,C**), suggesting that the unknown secreted factor was either a heat-labile small molecule or had an essential proteinaceous component.

Evolved clone P5-C produces CRAϕ phage particles

To identify the inhibitory factor present in P5-C supernatant, we turned to genetic evidence: what mutations had this strain accumulated during the evolution experiment? Sequencing the genomes of P5-C and other non-inhibitory clones from the same P5 population revealed several mutations in a putative prophage region that were unique to P5-C. First, we observed a 12.1-fold increase in the read-depth coverage of this region relative to the rest of the genome (**Figure 3.2A**). Individual reads were found that spanned a new DNA sequence junction, not found in the reference genome, connecting the ends of this region to one another in a way that is consistent with either circularization

of prophage DNA or the existence of tandem duplications of this prophage in the P5-C genome (**Figure 3.2B**). Three point mutations and a 75 bp deletion were also present in a most copies of the prophage region genes in the DNA sample from clone P5-C (**Figure 3.2B**).

The organization of this ADP1 prophage region resembles a filamentous phage(145), with the greatest similarity in gene content and layout to the *Vibrio cholerae* CTXϕ phage (139, 152). We therefore hypothesized that the increased DNA copy number of this region and the inhibitory effect of clone 5-C supernatant were due to the production of infective phage particles. We found that a polyethylene glycol (PEG) precipitation procedure designed to purify phage particles yielded a visible pellet from P5-C cell-free supernatant but not from the supernatant of wild-type ADP1. High-throughput sequencing of DNA isolated from the precipitated particles showed that they were further enriched for reads mapping to the prophage relative to the original genomic DNA sample from P5-C cells, with 84.9-fold increased read-depth coverage of the phage region versus the rest of the genome. Filamentous phages have circular single-stranded DNA genomes, so all of our results are consistent with clone 5-C having evolved to produce intact viral particles.

Phage DNA that was replicated and released from cells—but not packaged—could also potentially spread through an ADP1 population because of it high natural competence. Unpackaged phage DNA, perhaps released by cell lysis, might also be capable of inhibiting transformation by competing for DNA uptake sites with the exogenous construct containing an antibiotic resistance cassette that is used to measure transformation efficiency. However, we would not expect naked DNA to be inactivated by heat, as we observed for the inhibitory factor found in the supernatant, suggesting that clone P5-C was truly producing an intact virus.

To establish that viral particles with a protein capsid were being produced by P-5C, we imaged the phage after further purification via a cesium chloride gradient (**Figure 3.3A**). We observed elongated structures that resemble the morphology of other filamentous phages. To confirm that these structures were derived from the ADP1 prophage region that was amplified in clone P5-C, we separated proteins in the purified phages by SDS-PAGE and identified the proteins in the major bands using mass spectrometry (**Figure 3.3B**). The most abundant protein bands at <10 kDa corresponded to the product of the prophage gene *ACIAD1855*, which is positioned in the prophage sequence where one could expect to find the major coat protein (pVIII).

We were unable to find conditions in which this phage (hereafter referred to as CRAϕ) formed visible plaques on lawns of ADP1 cells. This result is not unexpected for a filamentous phage, as these phages bud from the surfaces of live cells rather than lysing their hosts (145). For filamentous phage infections, there is typically still a noticeable fitness cost to cells that are actively infected due to the resource drain of viral replication. The reduction in growth rate of wild-type ADP1 after exposure to P5-C supernatant is consistent with this sort of fitness cost. Thus, we used growth inhibition of ADP1 as a proxy for CRAϕ infection in further experiments.

Organization and evolution of the CRAϕ genome

In order to understand the potential effects of the mutations that evolved in our experiment on CRAϕ, we annotated its genes by comparing it to CTXϕ (**Figure 3.2B**). The integrated CRAϕ genome is similar in genetic structure to the prophage form of CTXϕ that is present in the genomes of Type II *V. cholerae* strains at the El Tor insertion site (152). Like these CTXϕ prophages, the CRAϕ integrant is flanked by two copies of a

repetitive sequence (RS) region containing the *rst* genes that mediate phage replication (153). In CRAϕ, the upstream RS1 copy and the downstream RS2 copy have identical nucleotide sequences. The CRAϕ RS regions lack the anti-repressor *rstC* gene found in certain CTXϕ RS sequences (154).

The CTXϕ prophage encodes several *Vibrio cholerae* virulence factors. We found evidence that some of these genes are also present in the CRAϕ prophage in *A. baylyi* ADP1. CRAϕ encodes a protein with amino acid sequence similarity to the CTXϕ zonula occludens toxin (*zot*) (155, 156). On the basis of its genetic organization, CRAϕ also appears to harbor two other CTXϕ virulence factors: the accessory cholera enterotoxin (*ace*) (157) and core encoded pilin (*cep*) (158). Zot, Ace, and Cep are thought to function as phage pI, pVI, and pVIII proteins, respectively (**Figure 3.2B**) (139). The ADP1 CRAϕ prophage does not appear to encode genes with homology to the Cholera toxin genes (*ctxA* and *ctxB*), which are located downstream of *zot* in CTXϕ.

The mutations found in CRAϕ in P5-C are two non-synonymous single base pair changes in the *orfU* (*ACIAD1854*), a synonymous base change in the hypothetical *ACIAD1848* reading frame, and a 75 bp deletion overlapping *rstR* (*ACIAD1849*) and *ACIAD1848* in the upstream RS1 copy of the *rst* region (**Figure 3.2B**). OrfU encodes the pIII protein that facilitates filamentous phage binding to the type-IV pilus and other host cell receptors (145). Mutations in phage tail-fiber proteins, that likewise mediate host cell recognition and infection, are often found in laboratory evolution experiments when there is co-evolution between a lytic virus and its bacterial host (159–161). It is possible that the pIII (OrfU) mutations expand the host range of CRAϕ phage produced by the P5-C clone, such that it has an improved ability to infect ADP1 variants that evolved by 1000 generations in the evolution experiment. As the P5-

C evolved CRAϕ phage still inhibits growth and transformation of wild-type (ancestral) ADP1, these pIII mutations do not seem to have restricted its host range.

The 75 bp deletion in P5-C CRAϕ includes the first 16 bp of the RS1 copy of the *rstR* ORF (*ACIAD1849*) and all of the intergenic space between *rstR* and *ACIAD1848*. RstR is the phage repressor protein (162), but this disruption of this gene is not expected to affect CRAϕ production because the genomic copy of *rstR* (*ACIAD1860*) in the RS2 *rst* copy is still intact. Instead, the key impact of this mutation is probably in removing RS1 RstR operator sites that control expression of the divergent *rstA* promoter, as suggested by the regulatory architecture in the *V. cholerae* CTXϕ prophage (163). Loss of phage repressor binding sites in this manner is known to create "ultravirulent" variants of other temperate phages that are able to superinfect cells that would normally be protected by genomic copies of the repressor protein (164). Indeed, the synonymous point mutation in the hypothetical *ACIAD1848* reading frame is also positioned where it could alter the RS1 *rstA* promoter, regardless of whether *ACIAD1848* actually encodes a protein. Overall, the mutations in the P5-C ADP1 clone are consistent with its apparent role as an overproducer of an ultravirulent CRAϕ phage with an expanded host range.

Non-competent ADP1 mutants are resistant to CRAϕ

Given the significant inhibitory effect of CRAϕ on the growth rate of the ancestral *A. baylyi* ADP1 strain, we next investigated whether any bacterial mutations that occurred during our evolution experiment protected cells from infection by re-activated phage. We tested whether P5-C supernatant inhibited the growth of five strains (**Figure 3.4**), each constructed from wild-type ADP1 by deleting one gene that was commonly found to be inactivated by mutations in the evolution experiment (144). Deletion of *pilB*,

comC, or *barA* made ADP1 completely resistant to growth inhibition by CRAϕ in our assay. Loss of *pgi* conferred partial resistance. Loss of *ACIAD3148* (a gene of unknown function) had no effect on CRAϕ inhibition of growth.

Extracellular DNA uptake by ADP1 requires a competence apparatus that is related to a type IV pilus (141), the same kind of structure that is targeted by other filamentous phage (145). The complete resistance of the *pilB* and *comC* deletion strains to CRAϕ suggested that the competence apparatus itself might be the receptor for CRAϕ infection. These strains have greatly reduced competence: we have previously shown that IS-element insertions into *pilB* and *comC* reduce transformation frequencies by 100- and 1000-fold, respectively (144). We further tested whether CRAϕ was able to inhibit the growth of a strain deleted for the gene *comP*, which is one of the core pilins in the competence apparatus(58), but was not found to be mutated in our evolution experiment. Deletion of *comP* renders ADP1 completely non-competent (165). As expected, the *comP* deletion mutant was also completely resistant to growth inhibition by CRAϕ (**Figure 3.4**), providing further evidence that the competence apparatus serves as the phage receptor.

In addition to the competence apparatus, *A. baylyi* ADP1 also expresses a type IV thick pilus that is involved in twitching motility and a type I thin pilus involved in adhesion (166, 167). Strains defective in genes required for competence, including *comC* and *comP*, still produce both types of pili and are fully motile (58, 62). Thus, the type IV pilus-related proteins of the competence machinery are believed to function entirely independently from the thick type IV pili involved in twitching motility (34). We found that our *pilB*, *comC*, and *comP* deletion mutants spread in soft agar as quickly as wild-type ADP1 (data not shown). This lack of a motility defect confirms that resistance to CRAϕ infection is not coincident with a change in the type IV thick pili.

We reasoned that extracellular DNA might act as a competitive inhibitor of CRAϕ binding and protect cells from infection if DNA and CRAϕ utilize the same receptor. However, we found that the addition of even very high concentrations of *E. coli* genomic DNA (up to 40 μg/mL), which are expected to saturate the DNA uptake sites on the surface of ADP1 (143), did not noticeably alleviate growth inhibition by the P5-C supernatant. This result suggests that DNA binding is not able to completely block CRAϕ from accessing its receptor and gaining entry into the cell, perhaps because CRAϕ and DNA utilize distinct, non-overlapping binding sites.

The complete resistance to CRAϕ infection exhibited by the *barA* deletion strain was unexpected (**Figure 3.4**). BarA operates as a hybrid sensory histidine kinase with the UvrY response regulator. The function of this two-component system has not been characterized in *A. baylyi* ADP1, but in *E. coli* and other γ-proteobacteria it activates expression of Csr RNAs and thereby regulates carbon metabolism, motility, adhesion, biofilm formation, virulence genes, and other stress responses (168, 169) upon sensing acetate or formate in the environment (170). We found that deletion of *barA* reduced ADP1 transformability by ~10-fold. Thus, the BarA/UvrY system appears to be necessary for fully inducing expression of competence in *A. baylyi* ADP1, and deletion of *barA* confers resistance to infection by CRAϕ for this reason. Interestingly, loss of *barA* also gave rise to a hypermotile spreading phenotype on soft agar, further indicating that the thick pilus involved in twitching motility is unrelated to the mechanism of CRAϕ infection.

Deletion of the *pgi* gene made ADP1 partially resistant to growth inhibition by CRAϕ (**Figure 3.4**). The *pgi* gene encodes glucose-6-phosphate isomerase, which is involved in extracellular polysaccharide biosynthesis in this strain. Knockout of *pgi* is also known to cause ADP1 cells to aggregate (144). Changes to the cell surface

associated with this mutation, possibly related to the production of polysaccharides that mask the phage receptor (171, 172) or by decreasing the surface area accessible for phage adsorption, may lead to this partial resistance phenotype.

CRAΦ emergence and evidence for coevolution in two laboratory populations

To learn more about the relationship between the phage and competence in the context of the long-term evolution experiment, we assayed CRAϕ emergence and population transformability at 100-generation intervals in two populations (P3 and P5). P3 was included see if CRAϕ had also reactivated in other populations besides P5. To detect phage, we used a PCR assay specific for the recircularized CRAϕ episome (see Materials and Methods). While developing this assay, we found that wild-type ADP1 produced a faint PCR band, which may indicate that phage is stochastically activated in some cells in a wild-type ADP1 population. In line with this hypothesis, treatment with mitomycin-C, a DNA-damaging agent that often induces prophage activation, increased the strength of this PCR band that is specific to re-activated phage.

We detected CRAϕ emergence as early as 300 generations in population P5 and 500 generations in population P3 (**Figure 3.5**). The presence of the circularized phage episome correlates strongly with reduced population transformability before 600 generations in each population, which is consistent with infective CRAϕ phage inhibiting transformation. We previously observed that IS*1236* insertions in the ADP1 genome that reduce or completely abolish competence arise and eventually become dominant within P3 (144). The timing with which cells with mutations in the competence apparatus reach high frequency in the P3 population, between 600 and 800 generations, is coincident with phage levels dropping below the detection limit of the assay, before staging a comeback

at 900 generations. In the P5 population, similar dynamics occurred later in the experiment. Non-competent clones rose to detectable levels in the population by 1000 generations and phage also seemingly suffered a setback at 900-1000 generations. Overall, these dynamics are suggestive of a co-evolutionary arms race in which CRAϕ emerges and infects the population, resistant ADP1 arise and become dominant in the population, and then new varieties of CRAϕ evolve that are able to continue to exploit the evolved bacteria.

To look for signatures of an evolutionary arms race and better understand the genetic basis of prophage reactivation, we isolated CRAϕ-infected cells from archived samples of each population and examined their phage sequences. By screening for small bacterial colonies, we were able to isolate CRAϕ-infected ADP1 clones from population P3 at 500 and 1000 generations and from population P5 at 300, 400, and 1000 generations. We sequenced the regions in which we previously found to mutations in the P5-C clone in these new CRAϕ isolates: the RS repeats and the *orfU-ace-zot* region (**Table 3.1**). The RS region was amplified with three different sets of PCR primers to distinguish among mutations in the RS1 and RS2 genomic copies and in the RS region contained within circularized phage episomes.

Most CRAϕ phages in each population have a mutation in the replication region: either the 75 bp deletion overlapping the RS1 *rstR* and *ACIAD1848* that was previously found in P5-C in population P5 or a 54 bp deletion overlapping *ACIAD1850* in population P3. While the function of the *ACIAD1850* gene is unknown, the proximity of the 54 bp deletion in this gene to *rstR* suggests that it may also affect phage repression, and this could be a parallel example of the evolution of an activated CRAϕ phage in P3. We observed several cases where there was a putative activating mutation in the *rst* genes in the phage episome without a corresponding mutation in the genome. In these cases, we

apparently found ADP1 cells became CRAϕ overproducers after super-infection by evolved phage rather than chromosomal mutations.

The phage isolated at 1000 generations from a P3 clone did not have any mutations in the regions that we sequenced. Therefore, it must have had an independent evolutionary origin from earlier phages found in this population at 500 generations. In contrast, we found evidence of long-term co-existence and co-evolution of CRAϕ in P5. Here, phages present at 1000 generations were most likely descended from earlier phages detected at 300 and 400 generations, as these isolates all shared the 75 bp RS1 deletion. The pIII (*orfU*) mutations present in the P-5C clone isolated at 1000 generations were not found in earlier isolates, further evidence that they appeared in response to the evolution of resistant or partially resistant ADP1 as part of an arms race. Again, we found a clone that had no mutation in the regions we sequenced at 300 generations in P5. Infection of this ADP1 clone apparently resulted from spontaneous prophage activation, phage mutations outside of the sequenced regions, or mutations elsewhere in the ADP1 chromosome that caused this particular clone to become a CRAϕ producer.

CRAϕ-like prophage are present in pathogenic *Acinetobacter baumannii* strains

To determine if CRAϕ was specific to ADP1 or more broadly conserved, we searched for similar filamentous prophage sequences in the genomes of other Acinetobacter species. We did not detect a similar prophage in the genomes of any sequenced Acinetobacter strains that have been demonstrated to be naturally transformable(26). However, we found eight strains of *Acinetobacter baumannii* with matches to key genes (*rstA*, *zot*, *orfU*, and *ace*) that enabled us to infer the presence of a CRAϕ-like prophage (**Table 3.2**). In contrast to *A. baylyi* in which twitching motility and

competence rely on distinct pili, *A. baumanii* twitching motility and competence appear to operate via the same type IV pilus, as mutations that disrupt one activity also eliminate the other (173, 174). Thus, it is unclear whether these putative *A. baumanii* phages infect and operate in a way that is directly analogous to how CRAϕ functions in ADP1.

Most *A. baumannii* genomes contain most or all of the genes required for competence, but this trait appears to have been lost in parallel in many Acinetobacter lineages very recently (175). In particular, there is evidence that deletions of *comP* often arise in *A. baumannii* strains. Recurrent loss of this key competence gene is reminiscent of how reactivation of CRAϕ selected against cells that maintained competence in our evolution experiment; it suggests that loss of competence may evolve in nature in response to avoid infection by filamentous phages.

Interestingly, all of the CRAϕ-like *A. baumannii* prophages that we identified are located near genes related to pilus assembly. Several cases of prophage sequences interrupting competence genes in naturally transformable bacteria were highlighted recently (176). In that study, these events were taken as evidence that phage indirectly benefit from targeting these sites to knock out competence: it prevents further transformation of homologous DNA derived from a non-infected genome that could purge the prophage sequence from an infected genome. Our work offers an alternative explanation. If a phage—like CRAϕ—requires the competence pilus as a receptor, then variants that knock out competence when they integrate into the host genome could gain an advantage because this prevents superinfection by their competitors.

CRAϕ-like prophages and their RS integration remnants often exist in many tandem copies in *A. baumannii* genomes. Strain AB5075-UW, for instance, has six complete prophage copies, and strain AB0057 has four. The various copies of these integrated CRAϕ-like phages in AB5075-UW and AB0057 vary somewhat in their gene

content and orientation, suggesting that they are not genome assembly artifacts (which could have arisen from the inadvertent inclusion of circular phage genomes in the sequenced DNA samples). All of the *A. baumanii* CRAϕ-like prophages carry genes with sequences that are orthologous to the known *V. cholerae* CTXϕ pathogenicity factors *zot* and *ace*. As is the case in *V. cholerae*, these prophage-encoded genes may impact host colonization. This possibility warrants the further study of these prophage proteins in pathogenic *A. baumannii* strains, to determine if they play a role in virulence.

DISCUSSION

Few Acinetobacter phages have been studied experimentally (177–180), although prophage sequences are common in their genomes (175). Here, we report the discovery of CRAϕ, a filamentous Acinetobacter phage that inhibits transformation and growth of *A. baylyi* ADP1. Filamentous phages are known to utilize type IV pili as their cell-surface receptors (145), and CRAϕ appears to target the ADP1 pilus involved in natural competence. Loss of function mutations in ADP1 genes required for DNA uptake via the competence apparatus confer complete resistance to CRAϕ. Mutations affecting extracellular polysaccharide production that increase aggregation of ADP1 cells also result in partial phage resistance. In sum, much of the phenotypic instability previously observed in long-term ADP1 cultures may be driven by the evolution of resistance to CRAϕ infection after the re-emergence of this virus from its genome.

We were unable to unambiguously determine what mechanism links the observed reduction in competence with CRAϕ infection. As adding high concentrations of purified DNA were not found to inhibit phage infection, it is unlikely that phage and DNA directly compete for binding to the competence apparatus. One possibility is that CRAϕ

infection reduces the expression of competence pili, resulting in fewer competence assemblies per cell and reducing the overall amount of DNA uptake during infection. This situation would resemble how *E. coli* phage M13, which infects via the type IV pilus involved in F-plasmid conjugation, represses expression of this receptor to prevent superinfection (181). If this hypothesis is correct, then the genomically integrated copy of CRAϕ in wild-type ADP1 appears to be incapable of this inhibitory function.

Selection for resistance to costly (though not fatal) infection by CRAϕ creates an evolutionary pressure for ADP1 to lose competence. We observed this outcome in our evolution experiment. Natural populations could be under a similar selection pressure to acquire loss of competence mutations to evade CRAϕ-like phage infection. Alternatively, a selection pressure from CRAϕ-like phages could favor cells that tightly control the expression of their competence genes until beneficial circumstances arise in which recombination in a population outweighs the risky exposure to phage infection (such as under stressful conditions or when adapting to a new environment). More broadly, phage targeting of DNA uptake machinery could help to explain the rarity of constitutive expression of competence in naturally transformable species.

Due to its unusually high natural transformability, ADP1 has been proposed as a next-generation chassis for genome and metabolic engineering (24). The emergence of CRAϕ in multiple experimental populations suggests that it represents a key vulnerability to using ADP1 as a model organism. In order to prevent CRAϕ from essentially nullifying the defining trait of ADP1, namely its competence, it will need to be deleted from the genome. Thus, this study demonstrates how studying adaptive evolution can be an important tool for guiding strain improvement: learning how strains evolve in undesirable ways can lead to specific strategies to counteract them (132). Another aspect of laboratory evolution that it illustrates is the discovery of new aspects of microbiology

that may go unnoticed in experiments with shorter durations. The re-emergence of this filamentous phage suggests that similar phages may limit horizontal gene transfer by the uptake of environmental DNA in nature and that CRAϕ -like prophage may contribute virulence factors to *A. baumannii* in a way that is reminiscent of *V. cholera* CTXϕ.

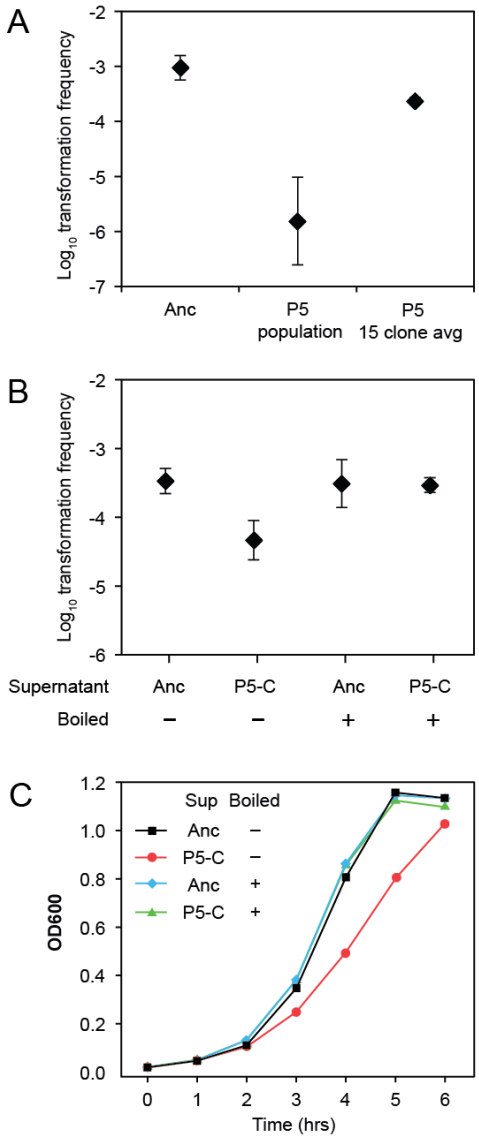

Figure 3.1. Laboratory-evolved *A. baylyi* ADP1 strain P5-C produces a factor that inhibits natural transformation and growth of wild-type ADP1 cells.

(A) Profiling the transformability of an evolved ADP1 population (designated P5) at 1000 generations unearthed the anomaly showing that transformation of the mixed population was significantly less efficient than expected from averaging measurements of 15 randomly selected clones tested independently. Error bars show 95% confidence intervals on log-transformed transformation frequencies. (B) Adding the cell-free supernatant from one of these fifteen clones (P5-C) inhibited transformation of wild-type ADP1, whereas cell-free supernatant from this ancestral ADP1 strain (Anc) had no effect on transformation. Boiling the P5-C supernatant ameliorated this inhibition. Error bars are as in panel A. (C) Cell-free supernatant from clone P5-C slows the growth rate of wild-type ADP1 cells. Again, cell-free supernatant from ancestral ADP1 had no effect on its own growth, and the inhibitory effect of P5-C supernatant was lost upon boiling.

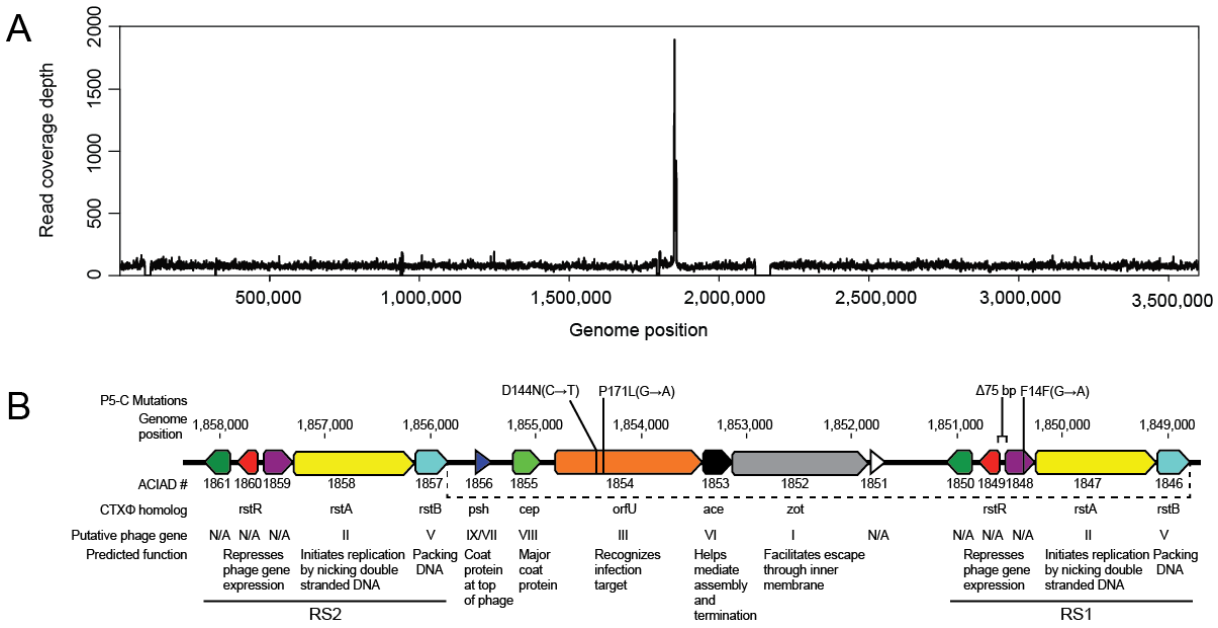

Figure 3.2. P5-C contains mutations in a putative filamentous prophage that is integrated into the *A. baylyi* ADP1 genome

(A) Illumina sequencing of DNA isolated from P5-C cells showed increased read-depth coverage of the prophage region relative to the remainder of the genome. Individual reads also spanned a new sequence junction that is consistent with the replication of circular ssDNA phage genomes by this strain as a result of reactivating the phage. (B) Genes in the prophage region are arranged in a manner that is consistent with it encoding a filamentous phage, which we have named the Competence Reducing Acinetobacter Phage (CRAϕ). We annotated the putative functions of CRAϕ genes based on homology to *V. cholerae* CTXϕ (139, 152) and other filamentous phages(145). Flanking copies of the repetitive sequence (RS) region containing the *rst* genes (*ACIAD1861-ACIAD1857* and *ACIAD1850-ACIAD1846*) have identical sequences in the wild-type ADP1 genome (indicated by shading in the same color). Mutations in this region in clone P5-C were localized to the RS1 copy. Circularization produces molecules consistent with the connection illustrated with a dashed line. Note that coordinates in this figure are displayed in a reverse orientation with respect to the reference genome.

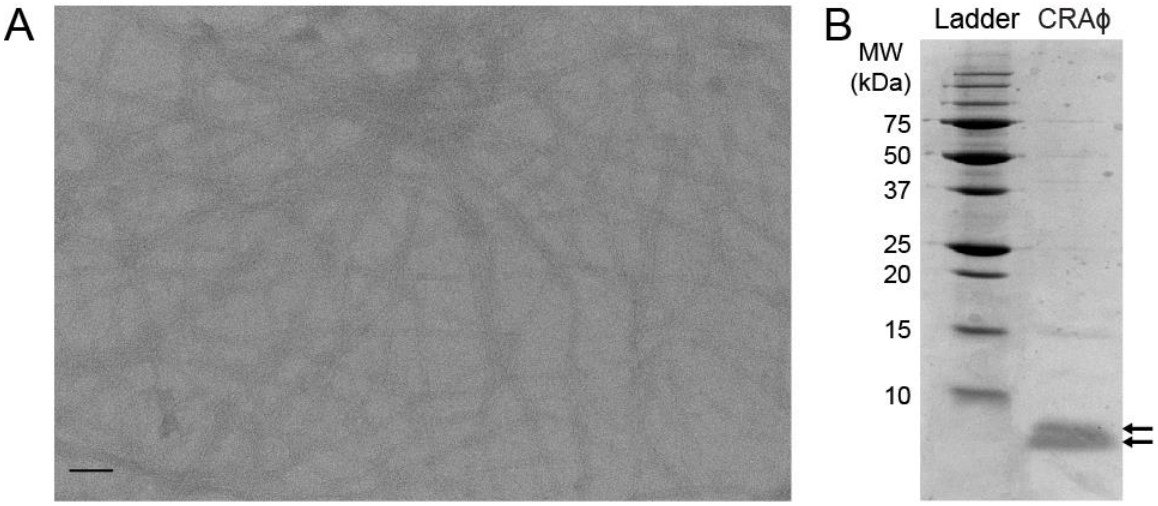

Figure 3.3. Visualization and biochemical analysis of CRAϕ particles.

(A) Transmission electron micrograph of purified CRAϕ shows elongated structures that resemble other filamentous phages. The scale bar is 50 nm. (B) SDS-PAGE gel of CRAϕ proteins. The two bands below 10 kDa (indicated by arrows) were excised and identified as being derived from the putative pVIII gene (*ACIAD1855*) by mass spectrometry.

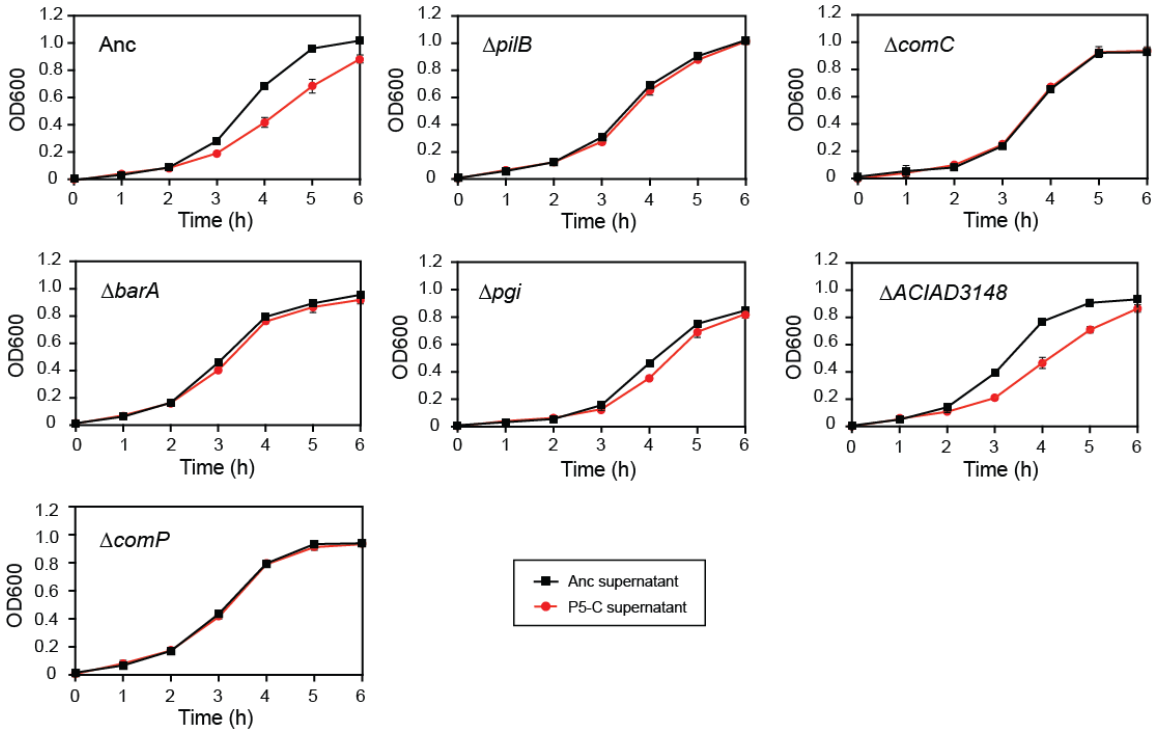

Figure 3.4. Knockout of ADP1 genes affecting competence confers resistance to CRAϕ.

Growth curves show the effect of sterile-filtered supernatant from the ancestral ADP1 strain (Anc) versus the CRAϕ-producing clone (P5-C) on the growth of various single-gene knockout mutants of wild-type ADP1. Loss-of-function mutations in *pilB*, *comC*, *barA*, *pgi*, and *ACIAD3148* were observed during the original evolution experiment in which reactivated CRAϕ emerged from the genome. The *comP* knockout was tested because this gene is known to be required for competence. Each strain-supernatant combination was tested in triplicate. Error bars show 95% confidence intervals on OD measurements, but they are often obscured by the symbols.

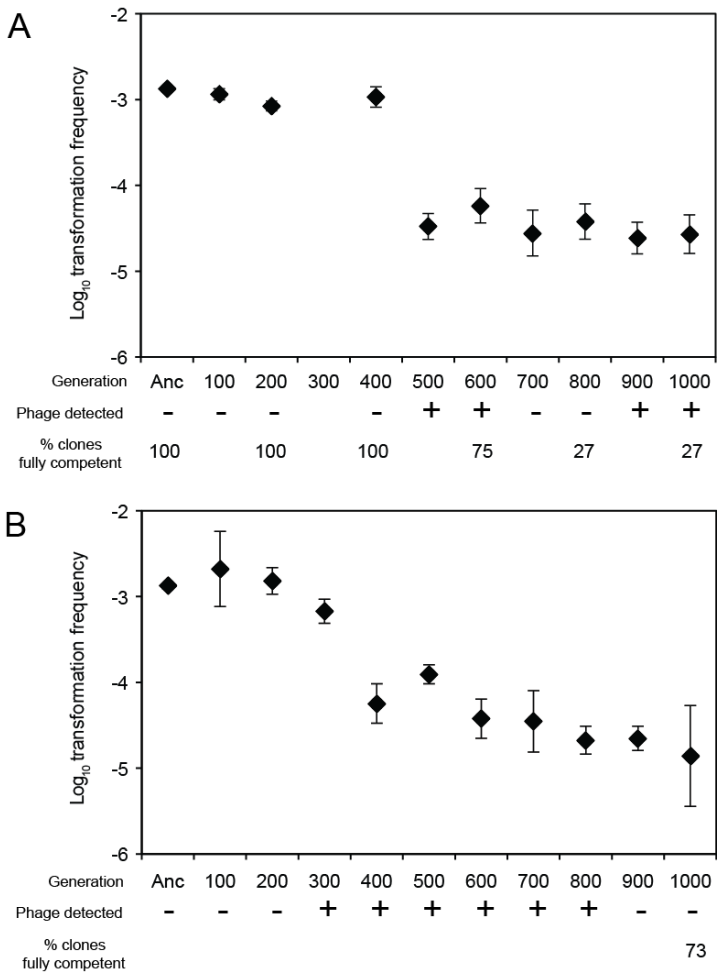

Figure 3.5. CRAϕ activation correlates with reduced population transformability and the evolution of reduced competence.

For two independent populations, P3 (A) and P5 (B), from a long-term evolution experiment(144), we assayed the overall transformability of cells in whole-population samples archived at 100-generation intervals (graphs). Error bars are 95% confidence intervals on log-transformed values. We further tested these samples for the presence of circularized CRAϕ DNA using a PCR assay (phage detected). For P3, data on the population structure of competence (% clones fully competent) at 200-generation intervals is from a prior study(144). The value for P5 at 1000 generations is from the current study. In each case, fully competent clones were defined as those having a transformation frequency of > 0.0002, and ≥10 clones selected randomly from the whole-population sample were tested at each time point. The sample for P3 at 300 generations was omitted because it apparently became contaminated with a CRAϕ phage derived from the one that was reactivated in clone P5-C during sample handling.

Population	Generation	Clone	Position[b]	Mutation	Genes affected
P3	500	A	(E) 1,850,875	Δ54 bp	*ACIAD1850*
P3	500	G	(E) 1,850,875	Δ54 bp	*ACIAD1850*
P3	1000	M	N/A	None	None
P5	300	B	(E) 1,850,538	Δ75 bp	*ACIAD1848*, *rstR*
P5	300	L	(UE) 1,850,538	Δ75 bp	*ACIAD1848*, *rstR*
P5	300	M	N/A	None	None
P5	400	M	(E) 1,850,538	Δ75 bp	*ACIAD1848*, *rstR*
P5	1000	C	1,850,343	G→A (P171L)	*orfU*
			1,850,395	C→T (D144N)	*orfU*
			(UE) 1,850,503	G→A (F14F)	*ACIAD1848*
			(UE) 1,850,538	Δ75 bp	*ACIAD1848*, *rstR*

Table 3.1. Mutations found in CRAϕ-producing clones from P3 and P5[a]

[a]Mutations in clone C from P5 at 1000 generations were identified by whole-genome sequencing. Sanger sequencing was used to identify mutations in the *rst* repeats and the *orfU-ace-zot* open reading frames of other clones. The regions sequenced are shown in Figure 3.2.
[b]Mutations found in the *rst* region were localized to either the upstream genomic (U) or episomal (E) copies, or both (UE). No mutations were found in the downstream genomic copy.

Strain	Genbank accession	Position[a]	Genes matched[b]	Organization[c]
AB030	NZ_CP009257.1	3,912,890	*zot*	r> ϕ> r>
AB031	NZ_CP009256.1	2,286,583	*zot*	r> ϕ> r>
AB0057	NC_011586.1	2,103,151	zot, ACIAD1848, *rstA* , *rstB*	ϕ> r> r> ϕ> r> ϕ> r> ϕ> r> r> ϕ> ϕ> r>
AB5075-UW	NZ_CP008706.1	1,989,509	*zot*, ACIAD1848, *rstA*, *rstB*	r> ϕ> r< ϕ< r< ϕ< r< ϕ< r< r< r< r< r< r<
AbH12O-A2	NZ_CP009534.1	1,913,110	*zot*, *rstA*	r> ϕ> r> ϕ< r<
AYE	NC_010410.1	1,899,034	*zot*, *rstR*	r< ϕ< r< ϕ< r<
6200	NZ_CP010397.1	3,481,874	*zot*, *rstA*	r> ϕ>
CIP70.10	NZ_LN865143.1	491,945	*zot*, *rstA*	ϕ< r<
D36	CP012952.1	1,990,179	*zot*, ACIAD1848, *rstA*, *rstB*	r< r< ϕ< r< r< r< ϕ< r<

Table 3.2. Putative CRAϕ-like prophages in *Acinetobacter baumannii* genomes

[a]Position of either the first phage gene or RS sequence.
[b]Genes with TBLASTN matches of $E < 10^{-5}$ when querying Acinetobacter complete genomes.
[b]Letters designate putative prophage (ϕ) and *rst* repetitive sequence (r) copies based on amino acid sequence matches to specific genes and conservation of gene organization. The orientation of each unit on the top (>) or bottom (<) genomic strand is also indicated.

Chapter 4: Factors affecting natural transformation rates in *Acinetobacter baylyi* ADP1

INTRODUCTION

Acinetobacter baylyi ADP1 is a naturally transformable Gram negative bacteria that has been proposed as an ideal model organism due to its high transformability and metabolic versatility (24, 25). A barrier to engineering in the strain however is that long stretches of genomic homology are required for efficient transformation of a construct into the ADP1 genome. Previous studies in the strain have shown that two major factors affect the transformation rate of a construct into genome are the length of genomic homology that flanks the construct and the concentration of that construct during a transformation (78, 95). As the length of flanking homology and DNA concentration increase, so does the transformation frequency up until a maximum transformation frequency of around 1% when using chromosomal DNA and concentrations ~2-4 µg/mL (0.84-1.7 pM) (95). Below 500 bp of flanking homology, transformation frequency decreases substantially (96), with transformations with 50 bp or less often being undetectable under normal transformation conditions (transformation frequencies < 10^{-8} – 10^{-9}). By understanding what factors may influence sub-500 bp flanking homology transformations, one can gain an overall better understanding limitations within the transformation process and perhaps identify strategies to improve sub-500 bp transformation frequencies. Finding paths towards optimizing the natural transformation system of ADP1 for improved transformation frequencies with shorter homologies would greatly improve its functionality for biological engineering and provide a novel tool for genome-scale engineering. In order to improve ADP1 transformation rates, however, understanding the mechanics of what limits transformation frequencies is required.

In order to provide a conceptual framework for our efforts to engineer greater transformation frequencies in ADP1, we first propose a general model to explain the overall relationship between homology length, DNA concentration, and transformation frequency in ADP1 (**Figure 4.1**). During the transformation process, exogenous double stranded DNA (dsDNA) is bound and translocated into the cytosol of the bacteria as single stranded DNA (ssDNA) in a 3′→5′ direction. As part of the DNA uptake process, part of the 3′ end of a given piece of DNA is cleaved to facilitate loading onto the machinery that translocates the DNA into the cytosol and converts it to ssDNA in the process. Given that ADP1 has no sequence specificity in regards to DNA binding (95), it can be assumed that the DNA translocation machinery binds to a random piece of the transforming DNA when initiating DNA uptake. In tests of linearized plasmid uptake in the homologous transformation system of *N. gonorrhoeae*, it was demonstrated that cleavage site of the incoming DNA was random (182). Assuming a similar process is present in ADP1, the consequence is that the internalized DNA pool has variable amounts of 3′ truncation relative to the original population of exogenous DNAs. In order for successful transformation, the internalized ssDNAs must have enough homology to the ADP1 genome remaining for a RecA-mediated homology search after RecA nucleation onto the ssDNA. The more homology those DNAs contain, the more efficient RecA-mediated recombination is (183) – resulting in improved transformation frequencies. The longer the flanking homology is on a piece of transforming DNA, the more likely it is that the binding and truncation location on this piece of DNA will leave a sufficient amount of homology present to allow recombination (i.e. the DNA is recombination proficient) with the ADP1 genome. In this manner, transformations with longer homology constructs will have a higher ratio of recombination-proficient internalized DNA fragments than those transformations with shorter homology. This increase in the

amount of internalized recombination proficient DNAs results in a higher transformation frequency. Increased DNA concentration during a transformation will also increase the amount of internalized recombination proficient DNA. While a higher concentration of transforming DNA will not change the relative amount of DNAs with a recombination-proficient amount of homology, it will increase the absolutely number of those DNAs proportionally with those with recombination-deficient homology lengths through an overall increase in the amount of internalized DNA. The overall higher number of longer segments subsequently increases the transformation frequency.

In this model, factors that affect the fate of internalized DNA may be expected to modify transformation frequencies. Specifically, competition for transformed ssDNA between ssDNA-specific nucleases and recombinases could limit recombinase access or degrade the regions of genomic homology present in the transformed DNA, reducing transformation rates. Such limitations are likely especially important when the DNA pool has a low about of homology to begin with. We tested the effects on transformation rates of several treatments that we predicted would modify the fates of internalized transforming DNA in ADP1. Our results suggest factors that affect the internalized DNA pool are important as nuclease and recombinase activity affected transformation rates. The improvements in transformation rate we were able to achieve by altering nuclease digestion of the transformed DNA and engineering recombination were however small and dependent upon the homology length and concentration of the construct.

MATERIALS & METHODS

Strains and growth conditions

The *Acinetobacter baylyi* ADP1 strain used for this work was the gift of Valérie de Crécy-Lagard (University of Florida). Unless otherwise noted, ADP1 was cultured at 30°C on the Miller formulation of Lysogeny Broth (LB) containing 10 g/liter NaCl. Liquid cultures were growth in 50 mL Erlenmeyer flasks or 18-mm by 150-mm test tubes with orbital shaking at 140 r.p.m. over a 1-inch diameter. Solid LB media included 1.5% (w/v) agar. Where appropriate, media were supplemented with 50 μg/mL kanamycin or 60 μg/mL spectinomycin. Frozen stocks of strains and evolved populations were stored at -80°C with added glycerol (20% v/v) as a cryoprotectant.

Assessing rates of natural transformation

In general, transformation assays were carried out by combining 1 mL LB with 70 μL overnight grown ADP1 culture with the appropriate amount of DNA in sterile glass test tubes and growing them at 30ºC and 200 r.p.m. overnight in a shaking incubator. These cultures were then diluted in sterile saline (0.85% w/v) and plated on selective and non-selective LB-agar plates using sterile glass plating beads. Plates were grown overnight at 30ºC before colonies were counted. Transformation frequencies were calculated as a ratio of CFUs on selective plates to CFUs on non-selective plates. When comparing effects across multiple homologies, equimolar amounts of transforming DNA were used. DNA used for transformations was either purified genomic DNA or PCR product prepared from genomic DNA. Unless stated otherwise, 1.5 nM concentrations of PCR products were used in all transformations and 2-3 replicates of each transformation assay were conducted.

For plate based transformations, 70 μL of overnight grown culture was mixed with an appropriate amount of DNA and then spotted onto 13 mm cellulose membrane

filters with a 0.025 μm pore size (Millipore) on LB-agar plates. These were allowed to dry then incubated overnight at 30ºC. After incubation, the filters were transferred into 1 mL sterile saline in 1.7 mL Eppendorf tubes, vortexed to suspend the cells in the saline solution, and plated with the appropriate dilutions to assay transformation frequency.

The preparation and assembly of transforming DNA

In this study, transformations were conducted either using genomic DNA prepared from an ADP1 strain containing a spectinomycin resistance gene (AB-*spec*^{R}) (amplified from pIM1463 (104) or from PCR products containing the resistance gene. AB-*spec*^{R} was constructed by flanking *spec*^{R} with ~1 kb of homology to the ADP1 genome through overlap PCR and transforming it into ADP1 such that the *spec*^{R} sequence displaced part of the non-essential gene *ACIAD0135* between the genome coordinates 139832 and 140099 as per previously described methods (24). PCR products containing the *spec*^{R} gene with varying amounts of homology were amplified from AB-*spec*^{R} genomic DNA. Where appropriate, phosphorothioates were added to the primers used in these PCRs to introduce the modifications into the PCR products. Before transformation, PCRs were purified using ethanol precipitation or the GeneJet® PCR purification kit (ThermoFisher Scientific) and quantified on a Nanodrop-1000.

For adding non-homologous DNA around a PCR product, 1 kb segments from pACYC184 and *lacI*-T5-*gusA*-pBAV1K that contain partial ORFs but no promoter sequences were amplified via PCR and added to the 5′ and 3′ segment of the *spec*^{R}-containing PCR fragment, respectively, with overlap PCR.

Construction of expression vectors

All plasmids were constructed using one-step isothermal Gibson assembly as previously described (184). The vector backbone used was *lacI*-T5-*gusA*-pBAV1k, a gift of Ichiro Matsumura (Emory University), which allows inducible expression of a gene inserted in place of the *gusA* coding sequence (104). For the construction of *lacI*-T5-*dprA*-pBAV1K, primers were designed to amplify the *dprA* (*ACIAD0209*) coding sequence from the ADP1 genome with homology to the vector region flanking the *gusA* ORF and separately, the pBAV1K backbone *sans* the *gusA* ORF was amplified via PCR. After isothermal assembly, the entire reaction volume was transformed directly into ADP1. The plasmids *lacI*-T5-*gfp*-pBAV1k and *lacI*-T5-*recET*-pBAV1k were similarly constructed. The ORF for *gfp* was retrieved from pBAVK-T5-*gfp* (104) and the *A. baumannii recET* ORF was retrieved from a variant of pAT02 (185), a gift of Bryan Davies (The University of Texas at Austin). Plasmid constructions were verified by prepping plasmid DNA from ADP1 cultures carrying the vectors and sequencing the boundaries of the inserted ORFs with Sanger sequencing, conducted at the UT Austin GSAF.

RESULTS

Nuclease protection of transforming DNA

A prime potential limiting factor for sub-500 bp homology transformations into ADP1 is the degradation of the ssDNA that enters the cytosol of ADP1 by native exonucleases. Exonuclease activity could compete with recombinases for access to internalized ssDNA and reduce the chance of that DNA becoming integrated into the

genome. In ADP1, DNA enters the cytosol in a 3′ → 5′ direction (60), while RecA nucleation occurs in a 5′→3′ direction (186), which provides nuclease protection for transformed DNA (81). This potentially leaves the entering 3′ end vulnerable to nucleases as well as the 5′ end of the transforming DNA after it enters the cytosol prior to RecA nucleation. DprA, a DNA binding protein specific to naturally competent bacteria has been shown to bind to transformed ssDNA, provide nuclease protection, and potentiate the loading of RecA onto the DNA, but it does not have a reported binding directionality and its competitiveness for the incoming 3′ end is unclear (80). It has been shown that knocking out 3′ → 5′ exonuclease ExoX and 5′ → 3′ exonuclease RecJ improves the transformation rate of constructs with low amounts of homology (78), supporting the important role nucleases play in ADP1 transformation.

Seeking a way to gain the competence improvement conferred by reduced nuclease activity without having to utilize a mutant ADP1 strain, we hypothesized that protecting the ends of the transforming DNA from nuclease activity would increase the transformation rates of a given construct. Given our model of transformation (**Figure 4.1**), the 3′ end of the transformed ssDNA is likely lost to cleavage, leaving the 5′ end as the best candidate for nuclease protection. To test the effect of 5′-acting exonuclease protection on transformation frequency, we tested the effect of adding phosphorothioate bonds to the 5′ ends of transforming DNA, a modification known to inhibit exonuclease activity (187, 188). We first tested for an effect of adding variable numbers of terminal phosphorothioate bonds to a PCR product containing a spectinomycin-resistance gene (*spec*$^{R)}$ and flanked with 500 bp homology to the ADP1 genome. In transformations with 0.078 nM (100 ng/mL) of PCR product, we found a significant (p = 0.014, one-tailed Student's t-test with Bonferroni correction) increase in transformation frequencies of ~3-fold with 2 terminal phosphorothioates (**Figure 4.2A**). This effect was maintained in

constructs with more than 2 terminal phosphorothioates. These results supports the hypothesis that 5′ acting exonucleases can play a role in preventing some transformation events.

Given our model of the potential impact of exonucleases, it might be expected that DNAs with shorter homologies may experience a larger relative transformability increase than longer homologies. Shorter homologies will take a shorter amount of exposure time to cytosolic nucleases be reduced below the minimum length required for efficient RecA-mediated recombination, potentially increasing the impact of nuclease protection for the transformability of those constructs. We tested this hypothesis by comparing the effect of two 5′ phosphorothioate linkages on transforming DNA with varying homology lengths (50 bp, 100 bp, 250 bp, and 500 bp) transformed into ADP1 in equimolar concentrations **(Figure 4.2B)**. To our surprise, we found that only at 250 bp was there a significant (p = 0.007, one tailed Student's t-test with Bonferroni correction) effect of the phosphorothioates (~5-fold increase) and we were not able to replicate our results previously seen at 500 bp. One potential explanation for this is the difference in DNA concentrations used between the two tests – the first test was conducted at 0.078 nM (100 ng/mL) whereas the second test was conducted at 1.5 nM (1910 ng/mL). A larger concentration of transformed DNA may saturate nuclease activity in the cell and limit the benefit of nuclease protection.

The effect of heterologous recombinase on transformation rates

A prediction of the competitive model between nuclease and recombinase activity for limiting transformation rates in ADP1 is that increasing the efficiency of low homology genome recombination should improve transformation frequencies. Using

native recombination machinery, efficient recombination into the genome requires large amounts of homology and large quantities of DNA for detectable transformation (189). To overcome this requirement in other systems, adding highly efficient phage recombinases to aid native recombination systems has allowed for construct integration with short homology (97). Theoretically, the expression of a similar recombinase system in ADP1 could similarly increase recombination rates for internalized DNA with shorter homology, increasing the relative availability of recombination-proficient DNA segments and improve transformation rates.

To test this hypothesis, we utilized a *recET* homolog isolated from *Acinetobacter baumannii* that has been used to integrate markers into its chromosome (185). As it was shown that the λ- red system does not appear to function in *A. baumannii*, we predicted that a functional recombinase system from a closely related species like *A. baumannii* was more likely to have an effect in ADP1. The *recET* homolog was cloned into an ADP1 expression vector under control of an IPTG-inducible promoter (104) and its effect on transformation rates at varying homology lengths was tested at two different DNA concentrations, 0.15 nM and 1.5 nM. Transformation rates measured with and without induction of the vector (**Figure 4.3A**) found no significant differences in transformation rates between induced and un-induced cells at 50 bp, 100 bp, or 500 bp flanking homology lengths at 1.5 nM of transforming DNA. Interestingly, we did see a significant (p = 0.005, one-tailed Student's t-test with Bonferroni correction) increase of ~2-fold at 250 bp, the same homology length with a significant effect from phosphorothioate protection. For the assays conducted at 0.15 nM DNA concentrations, all conditions except the 500 bp were below our detection limit (data not shown).

To get a sense of whether or not we had exceeded the saturating concentration of DNA in our conditions, we compared the transformation frequency of $spec^R$-500 bp at

0.15 nM and 1.5 nM in on uninduced cultures. We found that the high DNA treatment had a 9.4-fold increase over the low DNA treatment (data not shown). Given that one would expect a constant per-molecule transformation rate below a saturating concentration (78), a ~10-fold increase in transformation frequency correlates well to a 10-fold increase in DNA concentration, suggesting that the 1.5 nM concentrations do not exceed saturating concentrations for PCR products.

We went on to test if there were potential synergistic effects of phosphorothioates in the presence of recombinase induction at low homologies (**Figure 4.3B**). Under *recET* induction, adding two 5′-phosphorothioates resulted in a significant (p = 0.00054, two-tailed Student's *t*-test with Bonferroni correction) increase in transformation frequency of ~10-fold at 100 bp, but had no effect at 50 bp. These results indicate that transforming constructs with different amounts of homology may be limited by different cellular processes.

The effect of the heterologous expression of competence-specific DNA binding protein DprA on transformation rates

Another factor that could affect the relative rates of recombination and nuclease degradation in ADP1 is the activity of a recombination mediating gene found ubiquitously across naturally transformable species, *dprA* (80). DprA is a protein which binds to ssDNA as it enters the cytosol, protects it from nuclease digestion, and promotes the loading of *recA* to facilitate recombination (190). In some naturally transformable species, *dprA* is required for competence and its loss renders the strain completely non-transformable (81). In ADP1, loss of *dprA* results in a 10-fold reduction in transformation frequency (34). We predicted that increasing the expression of *dprA* in ADP1 would result in greater protection of transformed DNA, facilitate improved loading of RecA,

and subsequently lead to a higher rate of recombination into the ADP1 genome. We tested this by expressing *dprA* heterologously on an inducible plasmid (104). Much to our surprise, we found that expression of *dprA* leads to a >1000-fold reduction in transformation frequency in ADP1 across all tested homology lengths (**Figure 4.4**). In parallel, we tested the induction of the same vector carrying *gfp* to control for the effect of vector induction, which could adversely affect transformation by diverting cellular resources away from transformation-related processes. Induction of *gfp* did not affect transformation, indicating that *dprA* expression results were not an artifact of vector induction (**Figure 4.4**). This data suggests that *dprA* expression levels are not a limiting factor in natural transformation in ADP1. Furthermore, the data indicates that *dprA* in ADP1 might function in a regulatory capacity similar to its homolog in *S. pneumoniae* where it has been found to bind and inhibit competence regulating proteins from activating competence gene expression (191).

The effect of terminal protecting DNA on transformation frequencies

Put in the context of a nuclease based explanation for transformation limitation, the observation that increased homology correlates with improved transformation frequency can be interpreted in two ways. On one hand, the increased length could increase the amount of DNA that can be degraded by cellular nucleases before the construct drops below a recombination-proficient amount of homology. On the other hand, the increased homology allows for a longer stretch of RecA-bound nucleofilament to initiate recombination. While they are not mutually exclusive, the first hypothesis predicts that is not the added homology length, but the overall length of DNA flanking a section of homology that allows large homologous pieces of DNA to transform at high

rates. We tested this by flanking two transforming constructs with either 50 bp or 250 bp homology to the ADP1 genome with an additional 1 kb of DNA without any genomic homology. No significant difference in transformation frequencies was observed from the additional flanking DNA (data not shown). As before, 1.5 nM DNA concentrations were used. This result indicates that the benefit of longer homology is likely not nuclease protection, but rather improved recombination rates. Alternatively, the terminal protecting DNA could lead to improved nuclease protection that would increase transformation rates but adding non-homologous DNA could conversely reduce transformation rates by interfering with genomic integration – resulting in no net effect.

The effect of plate based transformation on transformation frequency

As mentioned previously, absolute transformation rates in ADP1 correlate directly with DNA concentration and transformation frequencies will increase in ADP1 up to a saturating concentration (78). While this concentration is ~1-1.5 pM for genomic DNA, our results suggest it could be higher than 1.5 nM for PCR products As DNA concentrations increase, the size distribution of 3′ truncated DNA segments inside the cell is unlikely to change, but increasing the overall number of DNA segments will lead to a larger total population of DNA with recombination-proficient homology and subsequently increase transformation rates. Under this model, strategies to increase the relative concentration of DNA in a transformation should yield relatively high transformation frequencies. In the *A. baylyi* field, it generally known that transformations where bacteria and DNA are mixed and then spotted onto media plates for growth can yield the highest transformation frequencies (Phoebe Lostroh, *personal communications*). This type of transformation is called a ‘puddle transformation’. As the puddle of bacteria

and DNA evaporates after spotting, the bacteria are deposited onto the agar plate for growth with the DNA in a very high relative concentration. Transformants are recovered by removing the bacteria after growth, resuspending them in liquid, and plating onto selective media. We quantitatively tested the difference in plate-based and regular liquid transformations for a finite amount of DNA and a set amount of starting cell inoculum. As with our other transformations, we used 1.5 pmol of DNA, either mixed in with the culture spotted onto the media plate or in the 1 mL liquid transformation volume. We assayed the effect at different homology lengths, as the model of improvement would predict a relatively consistent effect at each homology range. Our results (**Figure 4.5**) show ~10-fold improvement in transformation frequency across all homology ranges and no large difference in effect size between different homology lengths, supporting our model and indicating that saturating DNA concentrations are likely above 1.5 nM for PCR products.

DISCUSSION

In this study, we investigated factors limiting transformation rates in *A. baylyi* ADP1. The investigation was organized around a general model where transformed DNA exists as a population of variably 3′ truncated ssDNAs and whose recombination rates are dependent upon their concentration and total length of homology. While such a model generally explains the positive correlation between homology length, DNA concentration, and transformation rates, it is possible that differences between our proposed model and biological reality of ADP1 could exist and obfuscate conclusions drawn from the results of this study. For example, if ADP1 has an affinity to transform DNA starting from the ends of constructs in contrast to other characterized systems, the

model would be inherently flawed. Still, the data gathered is generally consistent for the above model and the effects of treatments that alter the fate of the internalized DNA pool appear to result in small but significant changes in transformability. The fact that such effects appear to be dependent upon homology length and DNA concentration fits well into the overall narrative created by our model regarding the potential limiting factors of ADP1 transformation rates.

The observation that the effect of 5′ nuclease protection on transformation frequency is generally consistent with the observation that loss of *recJ*, a 5′ acting exonuclease in ADP1, increases transformation frequency for constructs with shorter homologies but has no effect on longer homologies (77, 78). While we didn't see the same amount of transformation frequency change as was previously seen in *recJ* mutants nor did we see the increase across the same homology range, the previous work by others used a single base change rather than the insertion of a full gene as readout for transformation (78), which could affect the results. The lack of effect of nuclease protection at 50 bp or 100 bp of flanking homology could be a product of 3′-truncation: the 3′ ends of the internalized ssDNA would likely have a higher relative amount of recombination-deficient homology lengths and improved stability at the 5′ end of the DNA may not be enough to improve the recombination deficient caused by the shortened 3′ ends. If one of the DNA's ends is degraded below a minimum recombinase binding length of homology, any recombination event would likely be a product of homology-facilitated illegitimate recombination, which occurs in ADP1 at 1/10,000th the rate of homologous recombination (192). As homology lengths increase, the increase in the relative proportion of DNAs with recombination proficient 3′ ends could shift the limiting factor for recombination to be the relative longevity of both ends in regards to exonuclease degradation. Above a certain threshold (possibly ~500 bp), it is likely that

the cytosolic nuclease activity is not sufficient to reduce the transformed DNA population below a length where native recombinases are maximally effective. That we saw *recET* homolog expression improve transformation at the same homology length as the phosphorothioates suggests both that recombinase activity is also a limiting factor in ADP1 transformation and that this limitation is constrained to the same homology considerations outlined above regarding nuclease protection. With a small homology window available for potential improvement of transformation rates by nuclease and recombinase activity modification, it is likely that other avenues will be needed to seek improvements in constructs with <250 bp of flanking homology.

The difference in the effect of 5′ phosphorothioates between our low DNA test and high DNA test is intriguing. At low concentrations, exonuclease activity could be such that a substantial portion of the transformed DNA experiences a reduction of recombination proficient homology lengths – resulting in a net reduction in successful recombination and transformation. Conversely at high concentrations of DNA, DNA uptake of recombination proficient DNAs could outpace nuclease activity and result minimal influence of the nucleases on transformation frequency. Our data suggests our transformation conditions of 1.5 nM were not above saturating levels, but gives hints at where that value might be. If we assume that the only basis of the increased transformation rates in puddle transformations are the increased DNA concentrations, accept the work of others that shows per molecule transformation rates are constant below saturating DNA concentrations (78), and recognize that the puddle transformation DNA concentration was likely orders of magnitude greater than a 10-fold increase over a comparable liquid transformation, then a saturating DNA concentration is likely around 15 nM for PCR products. Under this assumption, the 10-fold increase we observed between puddle and liquid transformations represents the achievement of a saturating

concentration – the increase would have been much higher if there were a strict linear relationship between DNA concentration and transformation rate indicative of sub-saturating values.

Our results from *dprA* overexpression were surprising – increasing the amount of a protein that is thought to protect DNA from nuclease activity and promote loading of RecA onto the DNA logically should yield some improvement on transformation frequency, at least at the 250 bp range we found to be influenced by nuclease protection and heterologous recombinase expression. The most likely explanation for the ~1,000-fold reduction in transformation frequency we observed is that *dprA* has a role in regulating competence via a mechanism similar to *S. pneuomiae*. In *S. pneumoniae*, DprA interacts with one of the master regulators of competence gene expression (ComE) and mediates a secession of the competence state (191). Currently, the molecular mechanism regulating competence induction in ADP1 remains unknown. If *dprA* is involved in competence regulation, then studies regarding the interaction partners *dprA* has in ADP1 might give insight into the effector proteins that regulate transformation in ADP1. Identification of such effector proteins could allow for global overexpression of competence genes and improvements in transformation rates. Increasing the number of competence apparatuses may increasing the rate at which transforming ssDNA enters the cytosol, mimicking an increase in DNA concentration.

The lack of effect of terminal protecting DNA on transformations was interesting as it appears contrary to a competitive model between nuclease activity and recombination activity. It is possible however that the presence of long stretches of non-homologous DNA could negatively impact the recombination rates of the homologous region, negating any benefit gained by nuclease protection. A key question to both answer this question and to validate the proposed model as a whole is to directly assaying

the composition of the transformed ssDNA population during a transformation. Using radiolabeled PCR products with fluorophore tagged 3′ ends, one might be able to quantify both the amount of cleavage that occurs from the 3′ end and the amount and rate of degradation of the transformed ssDNA. Such studies could quantitatively answer questions week addressed qualitatively and provide insight into the molecular mechanism of DNA uptake in ADP1, especially if conducted in the conditions used in this study. In total, it is likely that the main driving force behind sub-500 bp transformation frequency is driven by the inherent constraints of homology length and DNA concentration in the context of mechanistic function of ADP1's competence machinery – truncation of internalized DNA and the absolute numbers of DNA with recombination-proficient homology lengths. Further study into the molecular mechanism of ADP1's transformation system will provide additional insight into this process and into potential avenues to push past transformation limitations towards hyper-competence.

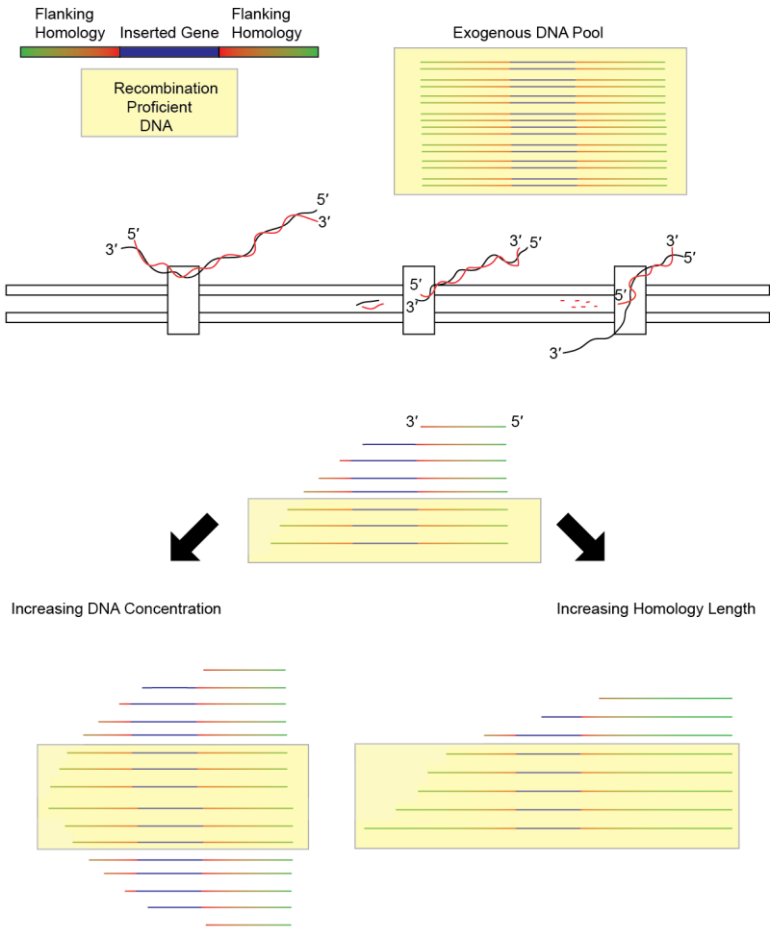

Figure 4.1. Model of the effect of homology and DNA concentration on internalized DNA pools.

The competence machinery of ADP1 randomly binds the exogenous dsDNA, converts it into ssDNA, and truncates it at the 3′ end. Once inside the cell, a proportion of that population will maintain an amount of the 3′ end sufficient to allow homologous recombination (yellow box). DNA with a recombination deficient length of homology (red) are unable to initiate genomic recombination. Increasing DNA concentration increased the absolute number of DNAs with recombination-proficient homology (green) and subsequently the recombination frequency into the genome (left). Increasing the length of homology increases the relative proportion of the recombination proficient DNAs and leads to increased recombination rates (right).

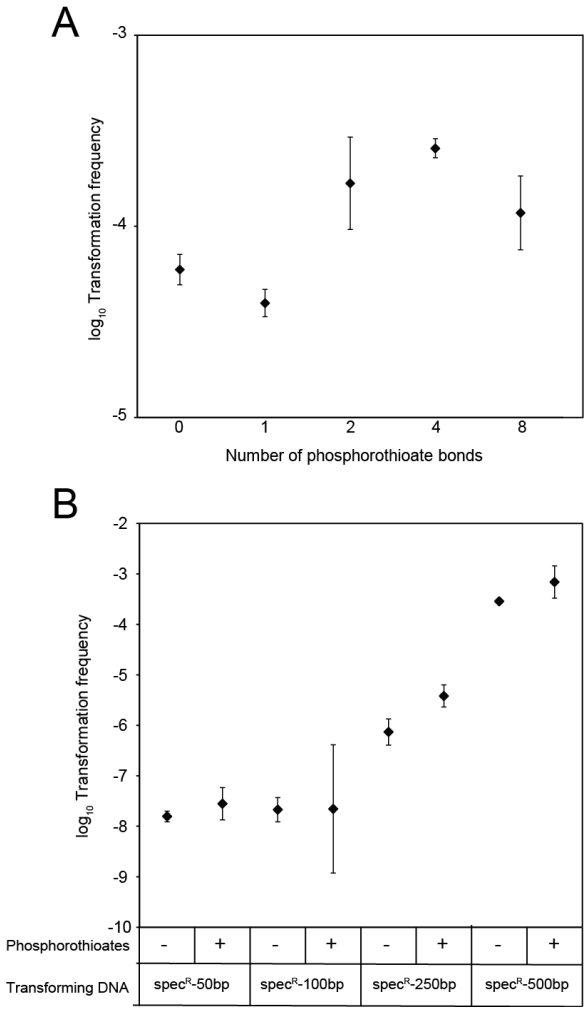

Figure 4.2. The effect of 5′ nuclease protection on transformation frequency.

(A) Variable numbers of phosphorothioate bonds were added to the 5′ end of a *spec*^{R} gene flanked with 500 bp of homology to the ADP1 genome and effected transformation frequency in a 0.078 nM transformation. (B) This effect of two 5′-phosphorothioate bonds was on transformation frequency also tested across variable homologies at a transforming DNA concentration of 1.5 nM. All experiments had $n = 3$ (except phosphorothioated *spec*^{R}-250 with an $n = 2$) and 95% confidence intervals are shown.

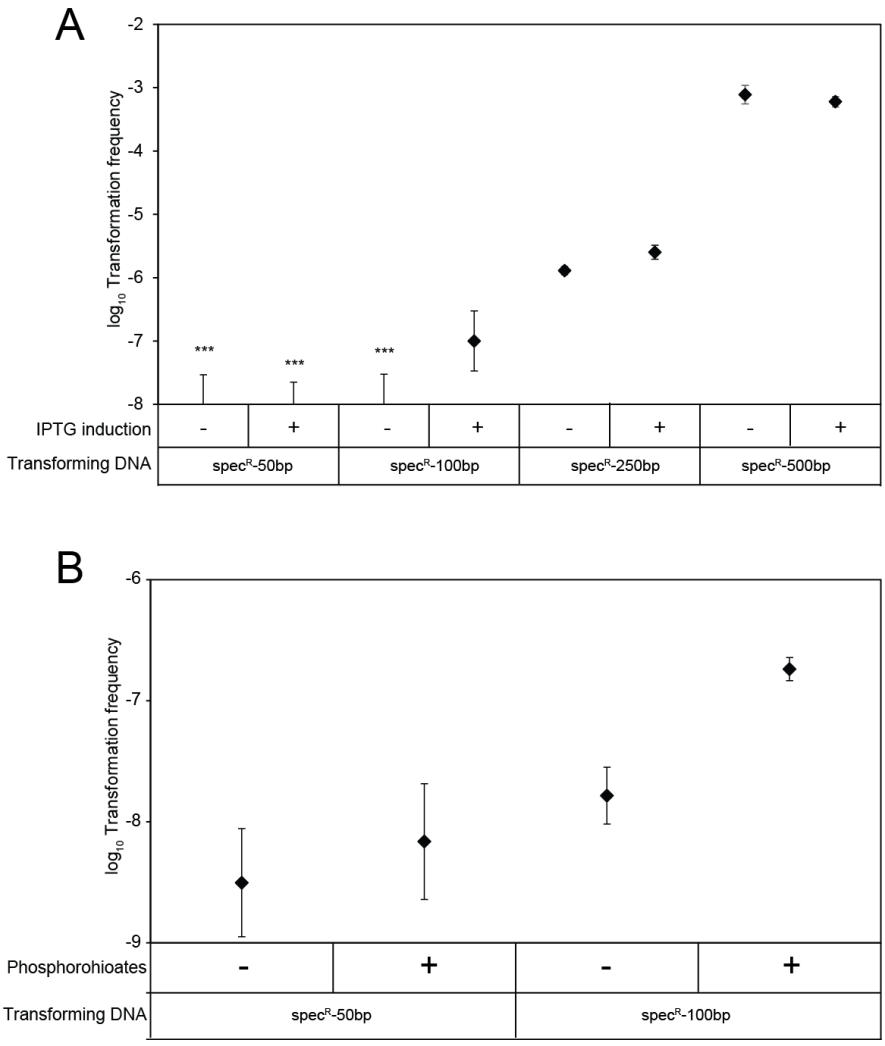

Figure 4.3. *A. baumannii recET* improves transformation rates at 250 bp.

(A) Induction of *recET* provides only marginally improvement at 250 bp. *** indicates that the results are below the level of detection of the transformation assay and the top error bar represents the detection limit based on the non-selective plating cell count. (B) 5′-phosphorothioated DNA combined with *recET* induction improves transformation frequency at 100 bp. $n = 3$ for all assays.

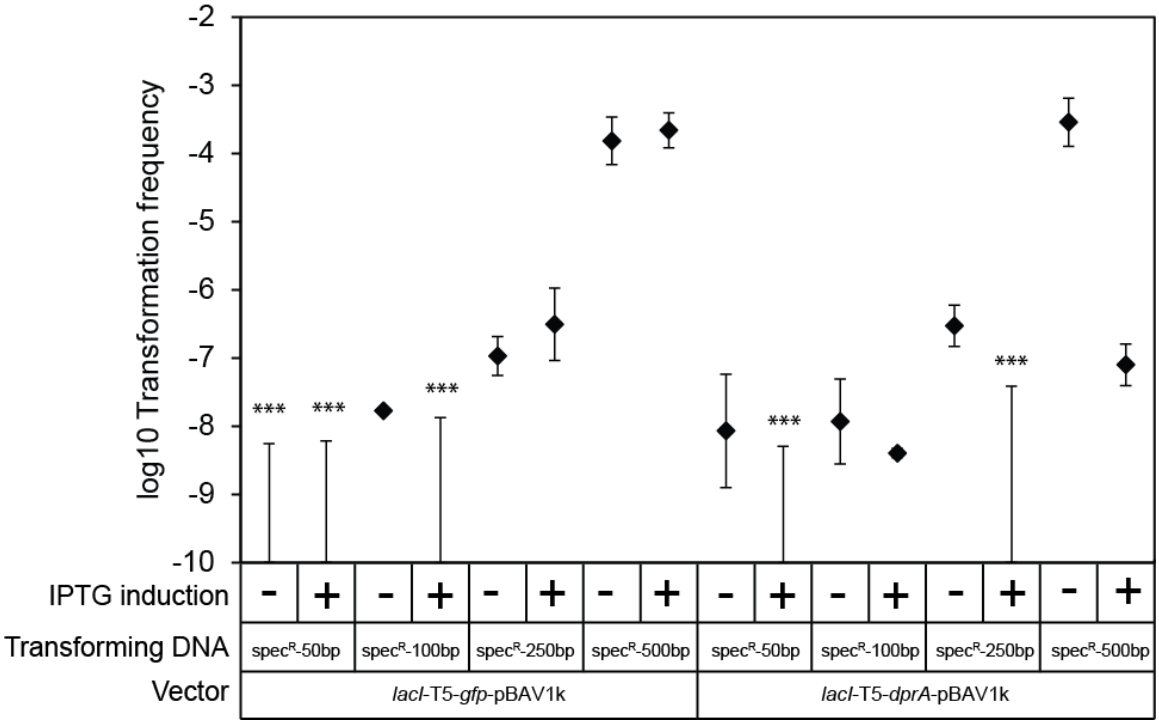

Figure 4.4. Heterologous *dprA* expression greatly reduces ADP1 transformability across multiple homology ranges.

Induced vs. uninduced ADP1 carrying vectors to express *dprA* or *gfp* were compared for the effect on transformation frequency across a range of homologies. *** indicates measurements were below the range of detection of the assay and the top error bar represents the detection limit based on the non-selective plating cell count. All experiments had $n = 3$ and 95% confidence intervals are shown.

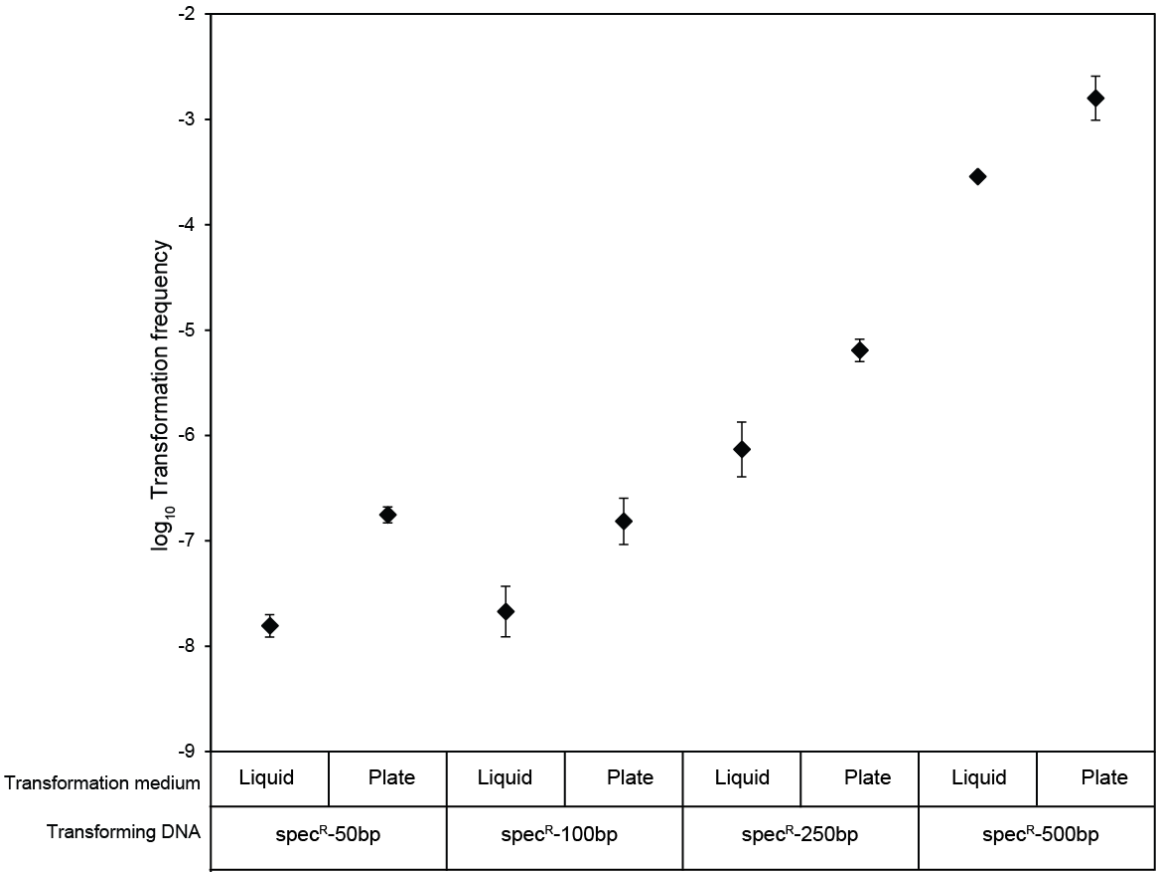

Figure 4.5. Plate-transformations increase transformation frequency across all homology ranges.

Plate based transformation increases transformation compared with liquid transformation given the same starting amount of DNA (1.5 pmol) and the same starting cell numbers. All experiments had $n = 3$ and error bars are 95% confidence intervals.

Chapter 5: Identification of genes required for *Acinetobacter baylyi* ADP1 competence using Tn-seq

INTRODUCTION

Horizontal transfer of genes into bacteria is an important process for evolution and for biotechnological applications. This is especially true for naturally transformable bacteria, which are a diverse set of bacteria that are able to uptake exogenous DNA and maintain it in a heritable fashion (12). Transformation in these bacteria is accomplished using a transmembrane competence apparatus, generally comprised of type IV pilins, that binds and translocates exogenous DNA into the cell. Characterizing the composition and regulation of this machinery is important for understanding the physiology of natural transformation and is vital for engineering this system for applications in biotechnology.

The mechanism and regulation of natural transformation has been studied in depth in a number of model systems (49). One species that has been less characterized in this regard is the Gram negative soil bacteria *Acinetobacter baylyi* ADP1. Proposed as an ideal model organism for biological engineering (24), at least 16 genes involved in competence have been identified and characterized (34), but it is unknown whether this list comprises the complete set of competence genes in ADP1. Furthermore, while it is known that competence in ADP1 develops after dilution from stationary phase into fresh media, the molecular mechanism of competence induction and the regulators of competence gene expression remain unknown.

Transposon insertion sequencing (Tn-seq) has emerged as an important tool in determining genotype-phenotype relationships on a genome-wide scale (193). A primary benefit of this approach is that the effect of a treatment can be assayed across a large number of randomly generated mutants simultaneously and quantified through next-generation sequencing. In this way, one can use selection to identify mutants of interest

on a genome-wide scale rather than having to individually generate and screen a large library of single gene knockouts. The current collection of known competence genes in ADP1 was identified by screening a mutant library constructed by transforming fragmented ADP1 genomic DNA ligated to an antibiotic selection marker (34, 165). As this study was conducted by manually screening picked mutants for competence effects, its genomic coverage is unknown. It is likely that there are competence-related genes in ADP1 that remain to be identified. Understanding what these genes are and how they might function could give key insights into natural transformation in ADP1.

In this study, we subjected a transposon-mutagenized ADP1 strain library to a transformation based selection scheme and used Tn-seq to conduct a genome-scale search for competence related genes. From our selection, we were able to identify at least two new competence involved genes: *pilR* and *ACIAD3188*. The identification of new competence genes from our data shows the utility of genome-scale approaches to defining the set of genes necessary for a complex trait and indicates that more factors than previously thought are involved in natural transformation in ADP1.

MATERIALS & METHODS

Strains and culture conditions

The *Acinetobacter baylyi* ADP1 strain used for this work was the gift of Valérie de Crécy-Lagard (University of Florida). Unless otherwise noted, ADP1 was cultured at 30°C on the Miller formulation of Lysogeny Broth (LB) containing 10 g/L NaCl. Liquid cultures were grown in 50 mL Erlenmeyer flasks or 18-mm by 150-mm test tubes with orbital shaking at 200 r.p.m. over a 1-inch diameter. Solid LB media included 1.5% (w/v)

agar. Where appropriate, media were supplemented with 50 μg/mL kanamycin (kan), 60 μg/mL spectinomycin (spec) and/or 5 μg/mL gentamicin (gent). *E. coli* in this study were grown at 37°C in LB and shaken at 225 r.p.m.

Construction of ADP1 transposon library

We constructed a transposon mutagenized library of ADP1 using suicide conjugation of the mini-transposon vector pBT20 (194). To set up the conjugation, 3 mL of *E. coli* β2163 (195) carrying pBT20 overnight culture and 1 mL of wild-type ADP1 overnight culture were pelleted via centrifugation and resuspended together in a single 100 μL volume of sterile saline (0.85% w/v). This mixture was then spotted onto a 13 mm cellulose membrane filter with a 0.025 μm pore size (Millipore) and placed on an LB-agar plate supplemented with 0.3 μM 2,6-diaminopimelic acid. After ~24 hours at 30˚C, the filter was removed from the plate and placed in a 1.7 mL Eppendorf tube with 1 mL sterile saline and vortexed to resuspend the bacteria. Of this mixture, 200 μL was removed to make dilutions for plating on LB-agar+gent to calculate the conjugation efficiency and the remaining 800 μL was used to inoculate 50 mL of LB+gent in a 250 mL Erlenmeyer flask. This culture was grown overnight in standard conditions. After growth, this culture was aliquoted into a number of individual 1.2 mL glycerol stocks with 20% glycerol (v/v) and stored at -80˚C.

Competence selection on ADP1 transposon library

DNA for transformation was prepared from strains of wild-type ADP1 carrying genes for kanamycin resistance (AB-kan^R) or spectinomycin resistance (AB-$spec^R$) integrated into the non-essential gene *ACIAD0135*, displacing the sequence between

genome coordinates 139832-140099 (Genbank: NC_005966.1). An entire ADP1 conjugation library glycerol stock containing ~ 10^9 cells was revived in 10 mL of LB+gent in a 50 mL flask that was then grown overnight in standard conditions. Before using this culture to start the selection experiment, the culture was preconditioned by pelleting 1 mL of the grown library, resuspending in 1 mL of LB, and using the entire volume to inoculate another 9 mL of LB+gent. This was again grown overnight in the same conditions.

To set up the first of round selections, 500 μL of the preconditioned library was added to 4.5 mL of LB+gent supplemented with the appropriate amount of DNA for the selection condition in a 50 mL flask and then grown overnight in standard conditions. For the control conditions, no DNA was added. Experiments in selection and control conditions were set up in triplicate. The unused preconditioned library was pelleted and frozen at -80˚C for later gDNA extraction and sequencing as our 'initial' library. For the first round of selections, the non-control transformations were pelleted via centrifugation, resuspended in 1 mL of LB, and then transferred into 49 mL of LB+gent+spec in 250 mL flasks. Prior to pelleting, 200 μL was removed for transformation assays. To simulate a transformation-like bottleneck, the no DNA control was diluted 1:10,000 in sterile saline and 1 mL of this was used to inoculate 49 mL of LB+gent. After overnight growth, the 250 mL cultures were used as the inoculum source for the second round of selection where kan selection is substituted for spec but otherwise unchanged from the conditions described above. After each round of selection, culture was pelleted and frozen at -80˚C for later gDNA extraction and sequencing.

Tn-seq library preparation

Libraries for Illumina sequencing were prepared as follows. Genomic DNA was extracted from populations stored at -80°C using a Purelink genomic DNA mini-kit (Thermo Fisher Scientific). Genomic DNA was sheared on a Covaris S220 to an average size of 300 base pairs. These fragments were poly-C tailed using terminal deoxynucleotidyl transferase (Promega) and then passed through size selection using AMPure XP© beads (Beckman Coulter) to eliminate very high and very low molecular weight library fragments. PCR with a biotinylated primer binding inside of the transposon and a non-biotinylated primer binding to the poly-C tail were then used to selectively amplify transposon-genome junction containing fragments. These fragments were then bound to streptatividin M-280 Dynabeads (Thermofisher Scientific), washed to remove unbound DNA, and then the remaining Illumina barcodes and primers were added in another PCR reaction. Internal barcodes were used to allow multiple libraries to be run with the same Illumina barcode. The libraries were pooled in equal ratios, analyzed on a BioAnalyzer 2100 (Agilent) for fragment distribution, and sequenced on an Illumina HiSeq 4000 at the UT Austin GSAF with a target of 125 million reads, or 5 million reads per library.

Computational analysis of Tn-seq data.

Scripts used to implement the Tn-seq analysis pipeline described in this section are freely available online (https://github.com/spleonard1/Tn-seq). Detailed usage instructions and descriptions are available at that location.

Identification of the locations of transposon insertions was conducted as follows. For each library, we searched each read for a 37 bp transposon specific junction sequence. Reads that contained this sequence were trimmed of the transposon containing

region and mapped to the ADP1 genome (Genbank: NC_005966.1) using Bowtie 2(196). We then tallied the number of reads that mapped to each insertion location, resulting in a single file for each library containing all identified insertion locations and the number of reads that mapped to each location. These counts were then used for analysis of changes in the abundance of transposon insertions between experimental samples.

To identify potential competence-involved genes, we compared the abundance of insertions in genes across experimental conditions. Count data for each library was binned by gene, and then analyzed using DESeq2 in comparison to the 'initial' pre-selection library. The standard DESeq2 workflow used includes normalization, dispersion estimation, and correction for multiple testing (197). For each gene a $\log_2$ fold change is calculated, which reflects the change in abundance between experimental conditions.

To remove genes with insertion numbers below what one would expect based on random distribution of the transposon throughout the genome – a method used to identify genes that may be essential, we employed a methodology used previously by others (198). Briefly, the insertion count file for the pre-selection library was LOESS smoothed to account for insertional bias near the chromosomal origin and then reads were binned by gene boundaries, giving a count of reads per gene. The total number of observed insertions was used to generate random datasets, each containing the same number of insertions but randomly distributed throughout the genome. The observed data was then compared to randomly generated data using DESeq2. Genes follow a bimodal distribution – either showing approximately the same number of insertions or far fewer than expected. Mclust was used to categorize ADP1 genes as either 'reduced' or 'unchanged'. The final list of essential genes contains only genes confidently assigned to the 'reduced' category (mclust uncertainty <0.01, padj <0.05). Genes predicted to be

'reduced' in this step cannot accurately be used for competence gene predictions and were excluded from our analysis.

Knockout strain construction and characterization

To generate DNA constructs used to create ADP1 knockout strains, we utilized a novel Golden Gate Assembly driven ADP1 transformation method. A derivative of the parts entry vector pYTK001 (199) carrying a selectable/anti-selectable knockout cassette (*tdk*-kan^R) flanked by Golden Gate BsaI compatible ends is used as the source of our selectable marker. To generate linear knockout constructs, ~1 kb regions flanking the target ADP1 genomic region are amplified by PCR with the addition of Golden Gate compatible BsaI sites with overhangs that allow annealing and assembly with the *tdk*-kan^R carried in the pYTK001 vector. Three pieces (the upstream PCR, the downstream PCR, and the *tdk*-kan^R carrying vector) are input into the Golden Gate reaction and the outcome is a linear construct of *tdk*-kan^R flanked by the appropriate homology to knock out a specific region. In general, knockouts were designed to retain ~5 codons at the beginning of the ORF and ~15 codons at the end to minimize any effects of the knockout on nearby genes. The entire volume of the Golden Gate reactions were added directly to standard transformations in liquid (1 mL LB + 70 µL overnight ADP1 culture) or on plates (50 µL overnight ADP1 culture spotted on LB-agar). Transformations were plated on LB+kan to isolate colonies that were then screened via PCR to verify successful gene knockout. Assays to determine transformation frequencies in knockouts were conducted with 100 ng/mL AB-*spec^R* gDNA as described in Chapter 3 of this dissertation. Assays to quantify twitching motility in these strains were performed as described elsewhere (174).

RESULTS

Selection for competence on a ADP1 Transposon library

In order to conduct a robust genome-wide screen for competence genes in ADP1, the traditional approach would have been to first knock out every non-essential gene individually and then assay each of those mutants for their competence vs. wild-type. While a single gene knockout collection of ADP1 exists (101), is it not currently distributed in its entirety, and performing replicated competence assays on each mutant in the collection would be prohibitive in its scale. A more efficient approach would be to use selection on a large mutant library to narrow down a list of genes to test for a potential effect on competence. With this in mind, we developed a transformation based selection approach that removes mutants defective in competence from a culture of ADP1 cells.

If one transforms a pool of mutants, any mutants that have lower transformation rates would be reduced in frequency after selection for successful transformants. In this way, one could identify genes required for competence by looking for a reduction in the frequency of individuals carrying knockouts of that gene after selection for transformation. To generate such a knockout mutant library and also enable the frequency of each mutant to be tracked, we used transposon mutagenesis (see Materials and Methods) and estimated the abundance of each mutant before and after selective transformation by using Tn-seq (200). By taking this approach, we can track global mutant abundance changes and identify potential competence involved genes without having to isolate and screen large numbers of clones from a post-selection library.

A diagram outlining the general selection scheme can be found in **Figure 5.1A**. We started with a transposon mutagenized ADP1 library and separated it into 3 selection

treatments with different transformation conditions. The DNA conditions used for the first round of transformation were as follows: 1 μg/mL *spec*^{R} + 500 bp homology linear PCR product for the most stringent transformation conditions (~10^{-4} transformation frequency), 100 ng/mL of *spec*^{R}-gDNA for moderately stringent transformation conditions (~10^{-3} transformation frequency), 2 μg/mL *spec*^{R}-gDNA for the least stringent transformation conditions (~10^{-2} transformation frequency), and no DNA was added for the control selection. All conditions were run in triplicate. Transformation conditions with different transformation frequencies were chosen to both capture a range of potential competence effects and minimize the risk of losing a mutant subpopulation to bottlenecks across the treatments. Two rounds of selection were conducted on our mutant library through the transformation of different antibiotic resistance genes into the same location in the ADP1 genome in each round. For a no transformation control, we diluted the mutant library to provide a similar bottleneck to transformation and otherwise grew the library in the same conditions but without transforming DNA or antibiotic selection. Transformation frequencies were tracked such that the bottlenecks for the various conditions could be monitored (**Figure 5.1B**). After two rounds of selection, Tn-seq libraries from each replicate at each round and for each condition (including the original preconditioned library) were prepared and sequenced.

Tn-seq on the starting pre-selection library showed 59,757 unique insertions covering 3073 of the 3325 (106) genes annotated in ADP1. Across our initial, selection, and control libraries, all had at least 4 million reads with the exception of replicate 3 of the 2 μg/mL round 2 population, which had less than 100,000 reads and was discarded from our analysis. We observed transposon insertions in 341 genes previously predicted to be essential in ADP1 (101), a difference that likely stems from our use of rich media rather than the minimal media used to construct the ADP1 knockout collection.

In order to measure the efficiency of our selection steps, we looked at the counts of transposon insertions into the region in gene *ACIAD0135* used to integrate selection cassettes in the ADP1 genome. As a transformation event should knock out any transposon integrated within this region by replacing it with the antibiotic resistance marker, successful selection should deplete these mutants from the population. Any transposons that map to this loci post-selection indicates non-selective carry over and allow us to examine the strength of selection. Comparing the read counts for transposon insertions in this region across our different conditions (**Figure 5.1C**), we can see that the first round of selection had very poor selection with only about 50% of the counts of transposons in this region lost after transformation compared to the no selection control. The second selection however was much more efficient, with 2-13% carryover depending upon the DNA condition.

Using PCA to compare the effect of our treatments and selection rounds on global changes in transposon abundance relative to the initial library, we can see that both rounds of selection tend to group all of the treatments with transforming DNA added together and away from the control treatments with no DNA (**Figure 5.1D**), suggesting that there is little effect of the different DNA treatments on the overall composition of our results. Therefore, we simplified our analysis of the selective and non-selective treatment by adding together the selective and non-selective $\log_2$ fold changes across treatments, rounds, and replicates, leaving us with one value for total selective $\log_2$ fold-change (TFC_S) and one for total control $\log_2$ fold-change (TFC_C). The more negative a TFC_S value is for a gene, the greater depletion under selective conditions. A negative TFC_C value on the other hand indicates that the gene is depleted from the population regardless of the imposed selection.

Known competence genes are depleted after selection for transformation

In order to determine if putative competence related genes were depleted by our selection, we first looked for known competence genes in our results. If the selection worked as intended, one would expect the distribution of TFC_S values for competence related genes to be lower than the TFC_S distribution for genes as a whole in ADP1. Comparing the distribution of TFC_S values across the subset of competence genes (**Figure 5.2B**) to the distribution of TFC_S across all genes (**Figure 5.2A**), we can see that competence genes congregate at a more negative TFC_S than the average gene, indicating that competence genes were preferentially depleted after selection. We did not find an association between the degree of depletion and the relative fold-change in competence loss expected from a gene knockout found in previous studies (i.e. gene knockouts that result in a complete loss of competence when deleted were not more depleted than knockouts that reduce transformability by only 10-fold when deleted) (**Table 5.1**). This might be expected, as the effects of knockout of competence genes in previous studies were not tested under our specific conditions. Noise introduced by differential fitness effects or random sampling through transformation mediated bottlenecking may also obfuscate any proportional effects of competence loss on mutant depletion.

Further evidence that the selection preferentially depleted competence genes can be found by ranking genes by TFC_S from lowest to highest and filtering out genes with a negative TFC_C. Mutants depleted in the control are likely to have a fitness cost associated with them, and this effect may result in more false positives for potential competence genes. This point is best illustrated by comparing the lowest 20 TFC_S genes with and without TFC_C filtering. Without filtering, the lowest 20 TFC_S list contains only one competence gene, *pilD*. After filtering for only positive TFC_C values, 7 out of 16 known competence genes in ADP1 appear in this top-20 list (**Table 5.2**). If such a list is

restructured to only consider the fold-change in the round 2 selection and not the less-selective round 1, then the number increases to 10 out of 16 known competence genes (data not shown). In summary, we find that our selection was successful in preferentially depleting known competence related genes.

Identification of new competence genes

Given that low TFC_S and positive TFC_C values appear to be generally indicative of competence genes, we hypothesized that genes with such values not previously demonstrated to be related to transformation could be involved in competence. Using the rank order low TFC_S positive TFC_C list we identified 10 potential targets (**Table 5.2**), the majority of which are putative or uncharacterized genes. To test if these genes might be involved in competence, we generated knockouts of five of these genes listed (**Table 5.3**) and assayed their competence at one of our experimental conditions, 100 ng/mL of $spec^R$ genomic DNA (**Figure 5.3A**) Knockouts of two genes significantly decreased in transformability: *pilR* (*ACIAD0258*) with 100-fold reduction in competence and *ACIAD3188* with a 1000-fold reduction in competence.

The *pilR* gene is part of a two-component transcriptional regulatory system. In *P. aeruginosa*, *pilR* controls the expression of type IV pilin structural gene *pilA* and regulates type IV pilin expression and twitching motility (201–203). To see if *pilR* might play a similar role in ADP1 and to query the potential effect of other knockouts on pilin expression, we assayed our mutants for twitching motility (**Figure 5.3B**). None of our mutants were deficient in twitching motility, indicating that *pilR* in ADP1 is not involved in regulating the expression of motility related pilin. While there appears to be a relationship between the natural transformation pilus and motility pilus in other bacterial

species, including in the closely related bacterium *A. baumannii* (201), there appears to be no overlap in ADP1 (34).

In *P. aeruginosa* as well as in ADP1, *pilR* is coexpressed in an operon with the sensor protein that facilitates its activation, *pilS*. Additionally, *pilR* works with *rpoN* (encoding σ^{54}) to activate transcription. It is possible that *pilR* works to activate competence-related pilin genes in ADP1 through a similar mechanism. If this is true, it stands to reason that *pilS* and *rpoN* mutants would have reduced competence and should be depleted in our data. We find that *pilS* and *rpoN* have higher TFC_S values than *pilR*, at 0.29 and -2.91 respectively. However, both have a higher TFC_C value than *pilR*, with values of 3.33 and 3.38 compared to *pilR*'s of 1.90. Furthermore, if we only look at round 2 data which had better competence selectivity, the TFC_S for *pilS* and *rpoN* drop to -2.44 and -3.65. This means that the combination of fitness-related enrichment of these genes and poor selection efficiency in round 1 could have countered the competence related depletion we observed in *pilR*. Ongoing testing will determine if *pilS* and *rpoN* have a role in ADP1 competence.

ACIAD3188, the other gene we found to be involved in ADP1 transformation, incurs a lot more uncertainty as to its specific function. Both *ACIAD3188* and the gene that is bicistronic with it, *ACIAD3189*, have sequences similar to genes found throughout the Acinetobacter genus but are not in other bacterial species. Although *ACIAD3188* does not contain any putative protein domains to give insight into its function, *ACIAD3189* contains an SdhE domain, indicating that the protein might be involved in succinate metabolism (204). In other species SdhE is bicistronic with a hypothetical membrane protein YgfX, a multimeric protein that associates with SdhE (205) and appears to interact with essential cytoskeletal proteins MreB and FtsZ (206). It is possible that loss of this protein in ADP1 affects cell division or the energy state of the cell in such a way

that loss of it blocks transformation, but further testing will be required to establish this connection.

DISCUSSION

In this study, we were able to successfully identify new competence related genes in *Acinetobacter baylyi* ADP1 by using transformation as a selective filter on a transposon mutant library and leveraging Tn-seq to track global changes in mutant abundance. Using the behavior of known competence genes within our selection scheme as a guide, we found that competence related genes generally had the greatest negative fold-change values under selective conditions but were not depleted under non-selective conditions. By sorting and ranking the genes in ADP1 by these criteria, we generated a list containing known competence genes along with genes that we predicted may be involved in competence due to their highly negative TFC_S and positive TFC_C values. Testing individual knockouts from this collection, we identified two new putative competence genes in ADP1: *pilR* and *ACIAD3188*. Work is ongoing to test additional mutants from this list and it is possible that additional competence genes will be identified as more knockouts are tested. It is also likely that our working list is incomplete. We filtered out competence genes that had a negative TFC_C value as the majority of known competence genes appear to have that feature. However, five (*pilD*, *comN*, *pilC*, *pilB*, and *comP*) appear to have a negative TFC_C value **(Table 5.1)**, indicating a possible fitness cost to their loss. With such mutants, it is difficult to differentiate between depletion that is caused by the fitness cost and depletion caused by the additional effect of reduced transformability.

The relationship between *pilR* and natural transformation in ADP1 is intriguing. Currently, the molecular mechanism for competence regulation is unclear, with only the observation that competence is optimally activated upon dilution into fresh media after reaching stationary phase. The few competence genes whose expression has been assayed relative to this process have shown that there is a disconnect between competence gene expression and competence induction, as these genes are induced in late-log phase when competence is reduced (58, 65). The role of *pilR* in controlling the expression of pilin genes is well documented and the lack of effect loss of *pilR* has on ADP1's twitching motility suggests that *pilR* may serve exclusively as part of the competence machinery in ADP1. Implication of *pilR* in ADP1 competence raises the question of whether *pilS*, the sensor that activates *pilR* (202), or σ^{54}, which acts in conjunction with *pilR* to activate pilin genes (207), might also be involved. Testing of the competence effects of *pilS* and *rpoN* knockouts are ongoing. If they affect ADP1 competence, it would represent the first potential molecular mechanism for the regulation of competence gene expression known in this species.

The role of *ACIAD3188* in natural transformation is more elusive. *ACIAD3188*'s location on an operon with *ACIAD3189*, putatively *sdhE*, tentatively identifies it as an ortholog to *ygfX*, which is highly conserved in *Enterobacteriaceae* (205). One key test could be to overexpress *ACIAD3188* and see if there is an effect on cell morphology, as YgfX has been shown to inhibit polymerization of cytoskeletal proteins (206). How such a protein may be related to transformation in ADP1 is unclear. It is possible that loss of *ACIAD3188* could alter the activity of *sdhE*, a flavination protein that is required for succinate dehydrogenase activity (204). Such a change may alter the energetics of the cell, something that has been proposed to be a factor in ADP1 competence induction (60).

In total, this work demonstrates the utility of genome-scale selection experiments and builds a framework to further our understanding of how natural transformation works in ADP1. These insights further our goal of engineering transformability in ADP1 to create a hypercompetent strain for use in genome engineering. Additionally, if all of the machinery required for transformation can be positively identified, it paves the way to consider transferring the complete set of competence machinery into non-transformable bacteria. By doing this, one could make any given bacteria naturally transformable and open up paths for whole new approaches for genome-scale engineering in bacteria.

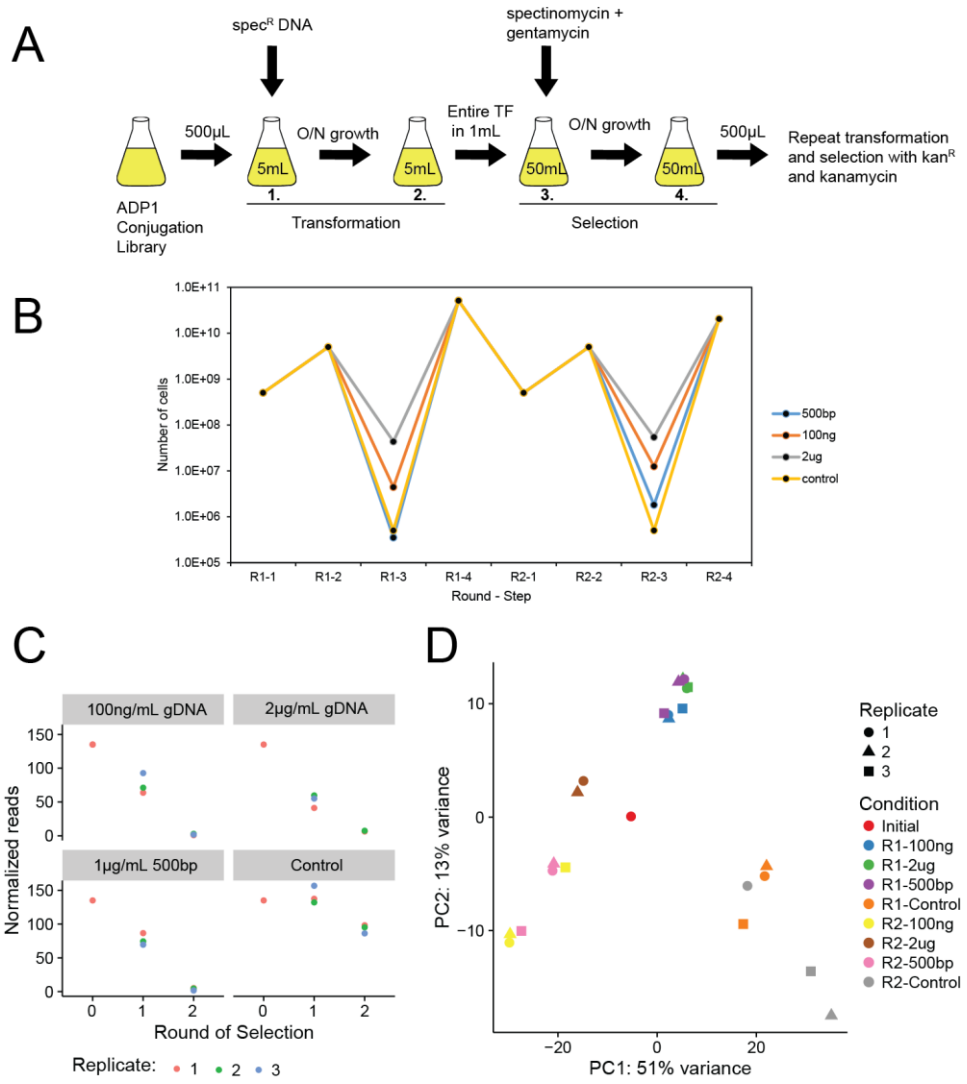

Figure 5.1. Transformation selection experiment.

(A) The selection scheme for the experiment. Two rounds of transformation and selection were conducted on all non-control populations, the first with spectinomycin selection and the second with kanamycin. Steps 1 and 2 comprise the transformation step and steps 3 and 4 comprise the selection step. (B) The population sizes present at each step in the selections, step 3 values were calculated based on the measured transformation frequency for that treatment during the selection experiment. The control was artificially bottlenecked at 10^{-4}. (C) Transposon insertions mapping to within the control region show poor selection efficiency after one round but greatly improved selection after two rounds in non-control populations. The control population demonstrates that the high rates of depletion in this region are not attributed to non-selective forced. (D) PCA of the $\log_2$ fold change of each population vs. the initial library shows strong divergence between control and treatments.

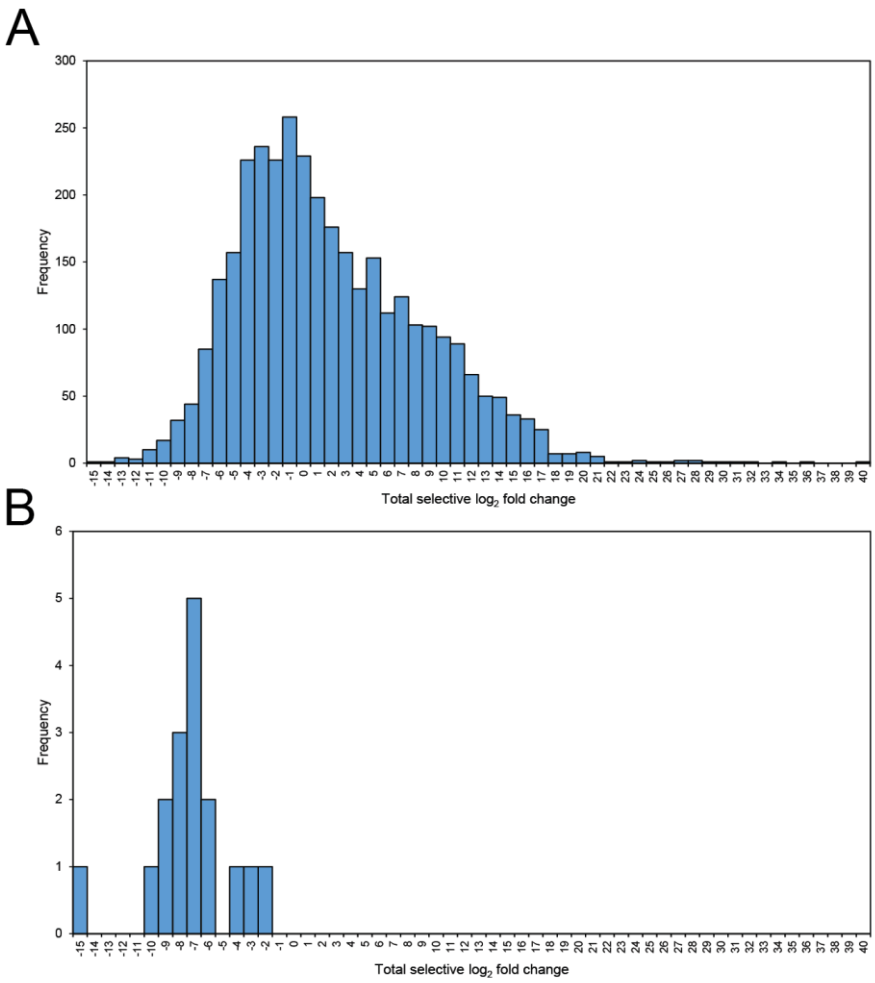

Figure 5.2. Selective depletion of competence genes.

Comparing the total selective $\log_2$ fold change (TFC_S) between all genes in ADP1 (A) and competence genes (B) shows a strong trend for competence genes to congregate in negative TFC_S space, supporting the efficacy of the selection in depleting competence genes relative to all genes.

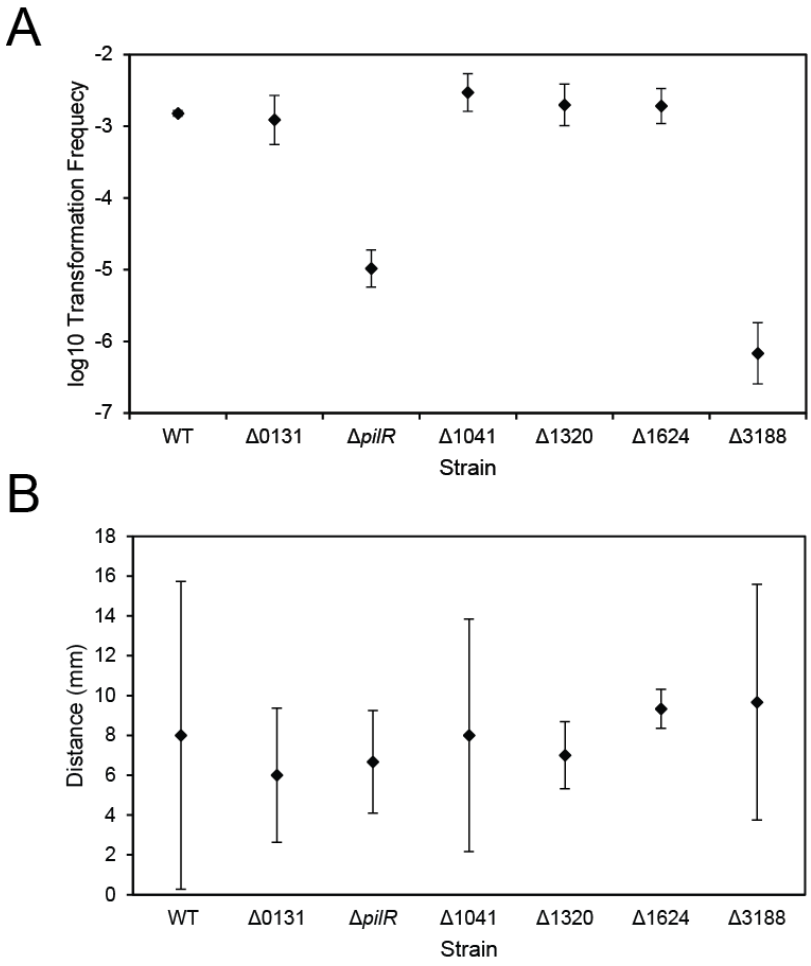

Figure 5.3. Reduced competence and unchanged motility in new putative competence genes.

(A) Transformation assays of potential competence genes show that knockouts of *pilR* and *ACIAD3188* reduce competence by 100- and 1000-fold respectively. (B) No twitching motility deficiency is observed relative to wild-type for any of the tested mutants. 95% confidence intervals are shown and $n = 3$ for all samples.

ID	Gene Name	TFC_S	TFC_C	Notes
ACIAD0360	*pilD*	-15.65	-2.28	KO: No data available
ACIAD0362	*pilB*	-10.07	-0.41	KO:100-fold competence reduction
ACIAD3064	*comEA*	-9.76	0.15	KO: Non-competent
ACIAD3318	*comB*	-9.13	0.80	KO: Non-competent
ACIAD3359	*comN*	-8.40	-1.08	KO: No data available
ACIAD3338	*comP*	-8.32	-0.41	KO: Non-competent
ACIAD3314	*comF*	-8.26	1.88	KO: 1000-fold competence reduction
ACIAD3355	*comQ*	-7.99	2.02	KO: No data available
ACIAD3360	*comM*	-7.98	2.47	KO: No data available
ACIAD3316	*comC*	-7.87	2.79	KO: Non-competent
ACIAD3356	*comL*	-7.85	3.21	KO: No data available
ACIAD0361	*pilC*	-7.13	-0.80	KO: No data available
ACIAD0209	*dprA*	-6.54	1.16	KO: 10-fold competence reduction
ACIAD2639	*comA*	-6.05	1.14	KO: Non-competent
ACIAD3357	*comO*	-3.63	0.66	KO: No data available
ACIAD3315	*comE*	-2.67	2.58	KO: 10-fold competence reduction

Table 5.1. Competence gene depletion in the Tn-Seq experiment

Data regarding the effect of competence gene knockout transformability gathered from elsewhere(34),

ID	Gene Name	TFC_S	TFC_C	Notes
ACIAD0135	NA	-10.74	0.90	Control region
ACIAD0131	**NA**	**-9.89**	**0.58**	**Tested in this study**
ACIAD3064	*comEA*	-9.76	0.15	Competence gene
ACIAD3318	*comB*	-9.13	0.80	Competence gene
ACIAD3188	**NA**	**-8.71**	**0.50**	**Tested in this study**
ACIAD3314	*comF*	-8.26	1.88	Competence gene
ACIAD3355	*comQ*	-7.99	2.02	Competence gene
ACIAD3360	*comM*	-7.98	2.47	Competence gene
ACIAD3316	*comC*	-7.87	2.79	Competence gene
ACIAD3356	*comL*	-7.85	3.21	Competence gene
ACIAD1320	**NA**	**-7.66**	**1.25**	**Tested in this study**
ACIAD2723	**NA**	**-7.34**	**0.03**	**Not tested**
ACIAD0717	**NA**	**-7.21**	**0.91**	**Not tested**
ACIAD1624	**NA**	**-7.16**	**0.04**	**Tested in this study**
ACIAD2586	**NA**	**-6.92**	**0.46**	**Not tested**
ACIAD2959	**NA**	**-6.78**	**0.52**	**Not tested**
ACIAD0258	***pilR***	**-6.60**	**1.90**	**Tested in this study**
ACIAD0209	*dprA*	-6.54	1.16	Competence gene
ACIAD1488	**NA**	**-6.45**	**0.82**	**Not tested**
ACIADtRNAAla_61	NA	-6.34	0.58	tRNA

Table 5.2. Lowest TFC_S gene hits with positive TFC_C values.

Bolded genes indicate new candidate competence genes and the lowest TFCS value gene, *ACIAD0135*, is assumed to be a result of transposon loss via selection marker integration.

ID	Gene Name	TFC_S	TFC_C	Notes
ACIAD0131	NA	-9.89	0.58	KO: No competence effect
ACIAD3188	NA	-8.71	0.50	KO: 1000-fold reduced competence
ACIAD1320	NA	-7.66	1.25	KO: No competence effect
ACIAD1624	NA	-7.16	0.04	KO: No competence effect
ACIAD0258	*pilR*	-6.60	1.90	KO: 100-fold reduced competence

Table 5.3. Tested candidate competence genes

Notes on competence measurements based on data presented in **Figure 5.3**.

Chapter 6: Conclusions and future directions

RESEARCH SUMMARY AND KEY FINDINGS

The overarching motivation of this work was to improve the naturally transformable bacterium *Acinetobacter baylyi* ADP1 as a model organism for biological engineering. A key advantage of ADP1 is its high recombinagenic potential due to its natural competence, which makes it easy to manipulate its genome. To improve the utility of this strain, two of its current limitations were addressed: the evolutionary instability of natural competence and the high homology requirement for efficient transformation of extracellular DNA into the genome. Learning about and solving these two problems would facilitate the creation of a strain that is highly transformable for low homology constructs and maintains this useful trait during growth in the lab.

To investigate the basis of the evolutionary instability of competence in ADP1, we conducted an adaptive long-term evolution (ALE) experiment (**see Chapter 2**). By characterizing losses of competence we observed in this experiment and finding the causal mutations using genome resequencing, we identified the genetic basis of competence loss and learned of a likely way to reduce its frequency. Predictably, competence loss was caused by mutations in genes known to be required for this trait. Interestingly, these mutations tended to be caused by IS*1236*, the sole mobile genetic element in ADP1. Taking a broader look at mutations in the ALE experiment, IS*1236*-mediated mutations represented an unexpectedly large proportion of mutations in evolved clones, sometimes more than 50% of all mutations. This indicates that IS*1236* is a primary overall mediator of inactivating mutations in ADP1. It is likely that removal of IS*1236* from the ADP1 genome would reduce the overall mutation rate and subsequently the rate of loss of competence mutants, improving the genetic stability of this phenotype. A mutation accumulation experiment also conducted on the strain, however, showed that

the spontaneous rate of loss of competence mutations in ADP1 is not high enough to explain the rate at which noncompetent sub-populations came to dominate in the ALE experiment. Loss of competence mutants must therefore provide a fitness benefit, and understanding why is vital for further improving the stability of this useful trait.

One critical reason for the fitness benefit conferred by loss-of-competence mutations was revealed by the discovery of the novel phage CRAφ that emerged during the ALE experiment (**see Chapter 3**). We found that this filamentous phage appears to be able to infect fully competent strains but is unable to infect strains carrying mutations in competence genes. A likely explanation for this is that filamentous phages tend to infect through type IV pili and this type of structure mediates the binding and translocation of DNA into ADP1 cells. While the CRAφ prophage is generally repressed in the ADP1 genome, we found that the appearance of mutations that likely inhibit phage synthesis correlated with the timing of CRAφ emergence. This suggests a model wherein mutant hyper-producers of CRAφ and/or supervirulent variants of CRAφ produced a strong selective pressure against bacteria with intact transformation machinery, thereby causing subpopulations with loss-of-competence mutations to come to prominence. Combining these results with the results from Chapter 2, there is a clear path forward to creating a strain with increased evolutionary stability of competence: removing both IS*1236* and CRAφ from the ADP1 genome. In this way, both a strong driver of loss of competence mutations and the fitness benefit of competence loss would be ameliorated.

To address the second problem in ADP1, the high homology requirement for efficient transformation, we hypothesized that manipulating processes that affect the fate of transforming DNA after it is internalized by the bacterium would lead to increased transformability of low homology constructs (**see Chapter 4**). By protecting the internalized DNA from nuclease activity and adding a phage recombinase, we sought to

increase the chance that a given piece of transformed DNA with a limited amount of homology would stay intact long enough to be recombined into the genome. We found little success with this approach, seeing only small improvements to transformability at intermediate flanking homology lengths (~250 bp) under specific conditions. The most promising improvement to transformation frequency at low homology lengths was from increasing the relative concentration of the transforming DNA via low volume transformations. It seems likely that this improvement stems from an overall increase of DNA flux into the cell, and it suggests that focusing on improving DNA flux could be the most fruitful approach to improving low homology transformation rates.

To improve DNA flux into ADP1, one approach might be to increase the number of active competence apparatuses per cell – a goal most directly achieved by increasing the expression of competence genes. Unfortunately, the mapping of competence genes in ADP1 is incomplete, and the molecular basis of competence gene regulation is as of yet unknown. Seeking to identify all of the competence genes and competence regulators in ADP1, we conducted a Tn-seq experiment and used transformation as a selective filter to identify these genes **(see Chapter 5**). From this experiment we were able to identify at least two new competence related genes, *pilR* and *ACIAD3188*. While the role of *ACIAD3188* is currently unknown, *pilR* is homologous to the response element in a two-component system that regulates pilin gene expression in *Pseudomonas aeruginosa*. As the ADP1 competence apparatus is a type IV pilus-like structure, is seems likely that this is the first competence regulator gene identified in ADP1. Work on testing potential competence genes identified by this experiment remains ongoing, and it is possible that additional competence genes will be found. Considering Chapters 4 and 5 together, we identified a likely path forward for improving ADP1 transformability – increasing DNA

flux into the cell – and found a key building block that could be manipulated to achieve that goal, *pilR* as the first putative regulator of competence gene expression in ADP1.

FUTURE DIRECTIONS – TOWARDS A HYPER-COMPETENT ACINETOBACTER STRAIN

While this work did not achieve the creation of a strain hyper-competent ADP1 strain that can transform low homology DNA efficiently, it did identify a path of inquiry that leads in that direction. This path follows two main trajectories: better understanding of the DNA uptake process and the regulation of competence genes.

The minimal success of the experiments in Chapter 4 to improve transformation rates shows that a better understanding the fate of internalized DNA is required in order to consider how manipulating that process can impact transformation rates. Transformations with radioactively labeled DNA could be used to track the size distribution and concentration of transformed DNA inside cells and test the relationships between internalized DNA pools and transformation rates. Furthermore, pulse-chase experiments using radioactive DNA followed by unlabeled DNA could be used to measure the degradation rate of internalized DNA as well as the physical rate of DNA uptake at different points during growth. With this data, a better model of DNA uptake can be built and revised hypotheses on how to improve transformation rates can be drawn. Along the same vein, we only tested a small number of modifications that effect the fate of intracellular DNA. Testing the effect of various knockouts of both nucleases and restriction systems on the fate of internalized DNA via radioactive DNA experiments and on transformation rates would prove informative as to the potential of this approach to elicit hyper-competence.

As little is known about ADP1 competence gene expression or regulation, a global characterization approach is required to first understand how the system behaves as a whole. Using RNA-seq to track competence gene expression levels during growth would provide this system wide view. With baseline wild-type expression levels, RNA-seq could also be used to test the effect of knockouts of putative competence regulators including *pilR* on competence gene expression to see which genes they control. Once regulators are identified, the effect of overexpressing them on transformation frequency can be tested.

An alternative approach to increasing competence gene expression is to bypass native regulation and create a refactored competence system by moving all competence involved genes to operons under control of a strong inducible promoters. Constructing a single consolidated artificial competence operon either in the ADP1 genome or carried on a vector could allow for optimization of the expression of each competence genes by combinatorial assembly of promoters and ribosome binding sites of varying strengths so that transformation rates could tuned to their optimal levels. Furthermore, different combinations of competence genes could be tested to see which genes are truly required for the apparatus assembly and which are involved in other competence related processes. By taking the refactored competence operon approach, one could develop a hyper-competent strain without a lengthy investigation into competence regulation in ADP1.

Altogether, this work has provided a number of key insights towards improving the ADP1 strain for use in biological engineering and has provided answers to a number of questions that have held back development in ADP1. With the foundation this work provides, future development of ADP1 as a model organism for biological engineering will lead to a new and improved tool for scientists to utilize as more and more ambitious genome editing projects are tackled.

References

1. **Griffith F**. 1928. The Significance of Pneumococcal Types. Journ Hyg **27**:113–159.
2. **Blount ZD**. 2015. The unexhausted potential of *E. coli*. Elife **4**:1–12.
3. **Mandel M**, **Higa A**. 1970. Calcium-dependent bacteriophage DNA infection. J Mol Biol **53**:159–162.
4. **Cohen SN**, **Chang AC**, **Hsu L**. 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Pnas **69**:2110–4.
5. **Datsenko K a**, **Wanner BL**. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A **97**:6640–6645.
6. **Wang HH**, **Isaacs FJ**, **Carr P a**, **Sun ZZ**, **Xu G**, **Forest CR**, **Church GM**. 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature **460**:894–898.
7. **Wenyan Jiang, David Bikard, David Cox, Feng Zhang and LAM**, **Jiang W**, **Bikard D**, **Cox D**, **Zhang F**, **Marraffini L a**. 2013. CRISPR-assisted editing of bacterial genomes. Nat biotechnol **31**:233–239.
8. **Huang CJ**, **Lin H**, **Yang XM**. 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. J Ind Microbiol Biotechnol **39**:383–399.
9. **Chen X**, **Zhou L**, **Tian K**, **Kumar A**, **Singh S**, **Prior BA**, **Wang Z**. 2013. Metabolic engineering of *Escherichia coli*: A sustainable industrial platform for bio-based chemical production. Biotechnol Adv **31**:1200–1223.
10. **Alper H**, **Stephanopoulos G**. 2009. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? Nat Rev Microbiol **7**:715–723.
11. **Beloqui A**, **de María PD**, **Golyshin PN**, **Ferrer M**. 2008. Recent trends in industrial microbiology. Curr Opin Microbiol **11**:240–248.
12. **Dubnau D**. 1999. DNA uptake in bacteria. Annu Rev Microbiol **53**:217–244.
13. **Johnsborg O**, **Eldholm V**, **Håvarstein LS**. 2007. Natural genetic transformation: prevalence, mechanisms and function. Res Microbiol **158**:767–778.
14. **van Dijl JM**, **Hecker M**. 2013. Bacillus subtilis: from soil bacterium to super-secreting cell factory. Microb Cell Fact **12**:3.

15. **Harwood CR**. 1992. Bacillus subtilis and its relatives: molecular biological and industrial workhorses. Trends Biotechnol **10**:247–256.

16. **Westers H**, **Dorenbos R**, **Van Dijl JM**, **Kabel J**, **Flanagan T**, **Devine KM**, **Jude F**, **Séror SJ**, **Beekman AC**, **Darmon E**, **Eschevins C**, **De Jong A**, **Bron S**, **Kuipers OP**, **Albertini AM**, **Antelmann H**, **Hecker M**, **Zamboni N**, **Sauer U**, **Bruand C**, **Ehrlich DS**, **Alonso JC**, **Salas M**, **Quax WJ**. 2003. Genome Engineering Reveals Large Dispensable Regions in Bacillus subtilis. Mol Biol Evol **20**:2076–2090.

17. **Radeck J**, **Kraft K**, **Bartels J**, **Cikovic T**, **Dürr F**, **Emenegger J**, **Kelterborn S**, **Sauer C**, **Fritz G**, **Gebhard S**, **Mascher T**. 2013. The Bacillus BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with Bacillus subtilis. J Biol Eng **7**:29.

18. **Brim H**, **McFarlan SC**, **Fredrickson JK**, **Minton KW**, **Zhai M**, **Wackett LP**, **Daly MJ**. 2000. Engineering Deinococcus radiodurans for metal remediation in radioactive mixed waste environments. Nat Biotechnol **18**:85–90.

19. **Lange CC**, **Wackett LP**, **Minton KW**, **Daly MJ**. 1998. Engineering a recombinant Deinococcus radiodurans for organopollutant degradation in radioactive mixed waste environments. Nat Biotechnol **16**:929–933.

20. **Gerber E**, **Bernard R**, **Castang S**, **Chabot N**, **Coze F**, **Dreux-Zigha A**, **Hauser E**, **Hivin P**, **Joseph P**, **Lazarelli C**, **Letellier G**, **Olive J**, **Leonetti J-P**. 2015. Deinococcus as new chassis for industrial biotechnology: biology, physiology and tools. J Appl Microbiol **119**:1–10.

21. **Blomqvist T**, **Steinmoen H**, **Håvarstein LS**. 2006. Natural genetic transformation: A novel tool for efficient genetic engineering of the dairy bacterium Streptococcus thermophilus. Appl Environ Microbiol **72**:6751–6756.

22. **Lindberg P**, **Park S**, **Melis A**. 2010. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. Metab Eng **12**:70–79.

23. **Elliott KT**, **Neidle EL**. 2011. Acinetobacter baylyi ADP1: Transforming the choice of model organism. IUBMB Life **63**:1075–1080.

24. **Metzgar D**, **Bacher JM**, **Pezo V**, **Reader J**, **Döring V**, **Schimmel P**, **Marlière P**, **de Crécy-Lagard V**. 2004. Acinetobacter sp. ADP1: An ideal model organism for genetic analysis and genome engineering. Nucleic Acids Res **32**:5780–5790.

25. **Young DM**, **Parke D**, **Ornston LN**. 2005. Opportunities for genetic investigation afforded by Acinetobacter baylyi, a nutritionally versatile bacterial species that is highly competent for natural transformation. Annu Rev Microbiol **59**:519–551.

26. **Vaneechoutte M**, **Young DM**, **Ornston LNN**, **De Baere T**, **Nemec A**, **Van Der Reijden T**, **Carr E**, **Tjernberg I**, **Dijkshoorn L**, **Baere T De**, **Nemec A**,

Reijden T Van Der, **Carr E**, **Tjernberg I**, **Dijkshoorn L**. 2006. Naturally transformable Acinetobacter sp. strain ADP1 belongs to the newly described species Acinetobacter baylyi. Appl Environ Microbiol **72**:932.

27. **Patel R**, **Mazumdar S**, **Ornston N**. 1975. beta-ketoadipate enol-lactone hydrolases I and II from Acinetobacter calcoaceticus. J Biol Chem **250**:6567–6577.

28. **Juni E**, **Janik A**. 1969. Transformation of *Acinetobacter calco-aceticus* (*Bacterium anitratum*). J Bacteriol **98**:281–288.

29. **Taylor WH**, **Juni E**. 1961. Pathways for biosynthesis of a bacterial capsular polysaccharide. J Biol Chem **230**:1231–1234.

30. **Park J**, **Park W**. 2011. Phenotypic and physiological changes in acinetobacter sp. strain DR1 with exogenous plasmid. Curr Microbiol **62**:249–254.

31. **Ramirez MS**, **Don M**, **Merkier AK**, **Bistué a. JS**, **Zorreguieta A**, **Centrón D**, **Tolmasky ME**. 2010. Naturally competent acinetobacter baumannii clinical isolate as a convenient model for genetic studies. J Clin Microbiol **48**:1488–1490.

32. **Touchon M**, **Cury J**, **Yoon E-JE-J**, **Krizova L**, **Cerqueira GC**, **Murphy C**, **Feldgarden M**, **Wortman J**, **Clermont D**, **Lambert T**, **Grillot-Courvalin C**, **Nemec A**, **Courvalin P**, **Rocha EPCP**. 2014. The Genomic Diversification of the Whole Acinetobacter Genus: Origins, Mechanisms, and Consequences. Genome Biol Evol **6**:2866–2882.

33. **Traglia GM**, **Quinn B**, **Schramm STJ**, **Soler Bistue A**, **Ramirez MS**. 2016. Serum albumin and Ca^{2+} are natural competence inducers in the human pathogen *Acinetobacter baumannii*. Antimicrob Agents Chemother AAC.00529–16.

34. **Averhoff B**, **Graf I**. 2008. The Natural Transformation System of *Acinetobacter baylyi* ADP1: A Unique DNA Transport Machinery, p. 119–139. *In* Gerischer, U (ed.), Acinetobacter Molecular Biology. Caister Academic Press, Norfolk, UK, UK.

35. **Huang WE**, **Wang H**, **Zheng H**, **Huang L**, **Singer a C**, **Thompson I**, **Whiteley a S**. 2005. Chromosomally located gene fusions constructed in Acinetobacter sp. ADP1 for the environmental detection of salicylate. Environ Microbiol **7**:1339–1348.

36. **Song Y**, **Li G**, **Thornton SF**, **Tompson IP**, **Banwart SA**, **Lerner DN**, **Huang WE**. 2009. Optimization of bacterial biosensors for toxicity assay of environmental samples. Environ Sci Technol **43**:7931–7938.

37. **Huang WE**, **Singer AC**, **Spiers AJ**, **Preston GM**, **Whiteley AS**. 2008. Characterizing the regulation of the Pu promoter in Acinetobacter baylyi ADP1. Environ Microbiol **10**:1668–1680.

38. **Rizzi A**, **Pontiroli A**, **Brusetti L**, **Borin S**, **Sorlini C**, **Abruzzese A**, **Sacchi GA**, **Vogel TM**, **Simonet P**, **Bazzicalupo M**, **Nielsen KM**, **Monier JM**, **Daffonchio D**. 2008. Strategy for in situ detection of natural transformation-based horizontal gene transfer events. Appl Environ Microbiol **74**:1250–1254.

39. **Gebhard F**, **Smalla K**. 1998. Transformation of Acinetobacter sp. strain BD413 by transgenic sugar beet DNA. Appl Environ Microbiol **64**:1550–1554.

40. **Ishige T**, **Tani A**, **Sakai Y**, **Kato N**. 2003. Wax ester production by bacteria. Curr Opin Microbiol **6**:244–250.

41. **Ishige T**, **Ishige T**, **Tani A**, **Tani A**, **Takabe K**, **Takabe K**, **Kawasaki K**, **Kawasaki K**, **Sakai Y**, **Sakai Y**, **Kato N**, **Kato N**. 2002. Wax Ester Production from n-Alkanes by Acinetobacter sp. Strain M-1: Ultrastructure of Cellular Inclusions and Role of Acyl Coenzyme A Reductase. Appl Environ Microbiol **68**:1192–1195.

42. **Santala S**, **Efimova E**, **Kivinen V**, **Larjo A**, **Aho T**, **Karp M**, **Santala V**. 2011. Improved triacylglycerol production in Acinetobacter baylyi ADP1 by metabolic engineering. Microb Cell Fact **10**:36.

43. **Santala S**, **Efimova E**, **Karp M**, **Santala V**. 2011. Real-time monitoring of intracellular wax ester metabolism. Microb Cell Fact **10**:75.

44. **Santala S**, **Efimova E**, **Koskinen P**, **Karp MT**, **Santala V**. 2014. Rewiring the wax ester production pathway of acinetobacter baylyi ADP1. ACS Synth Biol **3**:145–151.

45. **Kannisto M**, **Aho T**, **Karp M**, **Santala V**. 2014. Metabolic Engineering of Acinetobacter baylyi ADP1 for Improved Growth on Gluconate and Glucose. Appl Environ Microbiol **80**:7021–7027.

46. **Durot M**, **Le Fèvre F**, **de Berardinis V**, **Kreimeyer A**, **Vallenet D**, **Combe C**, **Smidtas S**, **Salanoubat M**, **Weissenbach J**, **Schachter V**. 2008. Iterative reconstruction of a global metabolic model of Acinetobacter baylyi ADP1 using high-throughput growth phenotype and gene essentiality data. BMC Syst Biol **2**:85.

47. **Kaplan N**, **Rosenberg E**. 1982. Exopolysaccharide distribution of and bioemulsifier production by Acinetobacter calcoaceticus BD4 and BD413. Appl Environ Microbiol **44**:1335–1341.

48. **Averhoff B**. 2004. DNA Transport and Natural Transformation in Mesophilic and Thermophilic Bacteria. J Bioenerg Biomembr **36**:25–33.

49. **Chen I**, **Dubnau D**. 2004. DNA uptake during bacterial transformation. Nat Rev Microbiol **2**:241–249.

50. **Stewart GJ**, **Carlson CA**. 1986. The biology of natural transformation. Annu Rev Microbiol **40**:211–235.

51. **Johnston C**, **Martin B**, **Fichant G**, **Polard P**, **Claverys J-P**. 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol **12**:181–96.

52. **Biswas GD**, **Sox T**, **Blackman E**, **Sparling PF**. 1977. Factors affecting genetic transformation of Neisseria gonorrhoeae. J Bacteriol **129**:983–992.

53. **Macfadyen LP**. 2000. Regulation of competence development in Haemophilus influenzae. J Theor Biol **207**:349–359.

54. **Magnuson R**, **Solomon J**, **Grossman AD**. 1994. Biochemical and genetic characterization of a competence pheromone from B. subtilis. Cell **77**:207–216.

55. **Meibom K**, **Blokesch M**, **Dolganov N**, **Cheng-yen W**, **School**. 2005. Chitin Induces Natural Competence in Vibrio cholerae. Science (80-) **310**:1824–1827.

56. **Charpentier X**, **Polard P**, **Claverys JP**. 2012. Induction of competence for genetic transformation by antibiotics: Convergent evolution of stress responses in distant bacterial species lacking SOS? Curr Opin Microbiol **15**:570–576.

57. **Palmen R**, **Buijsman P**, **Hellingwerf KJ**. 1994. Physiological regulation of competence induction for natural transformation in Acinetobacter calcoaceticus. Arch Microbiol **162**:344–351.

58. **Porstendörfer D**, **Gohl O**, **Mayer F**, **Averhoff B**. 2000. ComP, a pilin-like protein essential for natural competence in Acinetobacter sp. strain BD413: Regulation, modification, and cellular localization. J Bacteriol **182**:3673–3680.

59. **Herzberg C**, **Friedrich A**, **Averhoff B**. 2000. ComB, a novel competence gene required for natural transformation of Acinetobacter sp. BD413: Identification, characterization, and analysis of growth-phase-dependent regulation. Arch Microbiol **173**:220–228.

60. **Palmen R**, **Hellingwerf KJ**. 1997. Uptake and processing of DNA by Acinetobacter calcoaceticus - A review. Gene **192**:179–190.

61. **Bertrand JJ**, **West JT**, **Engel JN**. 2010. Genetic analysis of the regulation of type IV pilus function by the Chp chemosensory system of Pseudomonas aeruginosa. J Bacteriol **192**:994–1010.

62. **Link C**, **Eickernjager S**, **Porstendorfer D**, **Averhoff B**, **Eickernjäger S**, **Porstendörfer D**, **Averhoff B**. 1998. Identification and characterization of a novel competence gene, comC, required for DNA binding and uptake in Acinetobacter sp. strain BD413. J Bacteriol **180**:1592–1595.

63. **Porstendorfer D**, **Drotschmann U**, **Averhoff B**. 1997. A novel competence gene, comP, is essential for natural transformation of *Acinetobacter* sp. strain BD413. Appl Environ Microbiol **63**:4150–4157.

64. **Busch S**, **Rosenplänter C**, **Averhoff B**. 1999. Identification and characterization of comE and comF, two novel pilin-like competence factors involved in natural transformation of Acinetobacter sp. strain BD413. Appl Environ Microbiol **65**:4568–4574.

65. **Herzberg C**, **Friedrich A**, **Averhoff B**. 2000. ComB, a novel competence gene required for natural transformation of *Acinetobacter* sp. BD413: Identification, characterization, and analysis of growth-phase-dependent regulation. Arch Microbiol **173**:220–228.

66. **Lacks S**. 1962. Molecular fate of DNA in genetic transformation of Pneumococcus. J Mol Biol **5**:119–131.

67. **Lehman I**, **Nussbaum A**. 1964. The deoxyribonucleases of Escherichia coli. J Biol Chem **239**:2628–2636.

68. **Chase J**, **Richardson C**. 1974. Exonuclease VII of Escherichia coli. J Biol Chem **249**:4553–4561.

69. **Connelly JC**, **Leach DRF**. 1996. The sbcC and sbcD genes of Escherichia coli encode a nuclease involved in palindrome inviability and genetic recombination. Genes to Cells **1**:285–291.

70. **Viswanathan M**, **Lovett ST**. 1999. Exonuclease X of Escherichia coli. J Biol Chem **274**:30094–30100.

71. **Goldmark PJ**, **Linn S**. 1972. Purification and properties of the recBC DNase of E. coli K-12. J Biol Chem **247**:1849–1860.

72. **Viswanathan M**, **Lovett ST**. 1998. Single-strand DNA-specific exonucleases in Escherichia coli: Roles in repair and mutation avoidance. Genetics **149**:7–16.

73. **Kickstein E**, **Harms K**, **Wackernagel W**. 2007. Deletions of recBCD or recD influence genetic transformation differently and are lethal together with a recJ deletion in Acinetobacter baylyi. Microbiology **153**:2259–2270.

74. **Lovett ST**, **Clark AJ**. 1984. Genetic analysis of the recJ gene of Escherichia coli K-12. J Bacteriol **157**:190–196.

75. **Han ES**, **Cooper DL**, **Persky NS**, **Sutera VA**, **Whitaker RD**, **Montello ML**, **Lovett ST**. 2006. RecJ exonuclease: Substrates, products and interaction with SSB. Nucleic Acids Res **34**:1084–1091.

76. **Sharma R**, **Rao DN**. 2009. Orchestration of Haemophilus influenzae RecJ Exonuclease by Interaction with Single-Stranded DNA-Binding Protein. J Mol Biol **385**:1375–1396.

77. **Harms K**, **Schön V**, **Kickstein E**, **Wackernagel W**. 2007. The RecJ DNase strongly suppresses genomic integration of short but not long foreign DNA

fragments by homology-facilitated illegitimate recombination during transformation of Acinetobacter baylyi. Mol Microbiol **64**:691–702.

78. **Overballe-Petersen S**, **Harms K**, **Orlando L a a**, **Mayar JVM**, **Rasmussen S**, **Dahl TW**, **Rosing MT**, **Poole AM**, **Sicheritz-Ponten T**, **Brunak S**, **Inselmann S**, **de Vries J**, **Wackernagel W**, **Pybus OG**, **Nielsen R**, **Johnsen PJ**, **Nielsen KM**, **Willerslev E**, **Vries J De**. 2013. Bacterial natural transformation by highly fragmented and damaged DNA. Proc Natl Acad Sci U S A **110**:19860–5.

79. **Harms K**, **Wackernagel W**. 2008. The RecBCD and SbcCD DNases suppress homology-facilitated illegitimate recombination during natural transformation of Acinetobacter baylyi. Microbiology **154**:2437–2445.

80. **Mortier-Barrière I**, **Velten M**, **Dupaigne P**, **Mirouze N**, **Piétrement O**, **McGovern S**, **Fichant G**, **Martin B**, **Noirot P**, **Le Cam E**, **Polard P**, **Claverys JP**. 2007. A Key Presynaptic Role in Transformation for a Widespread Bacterial Protein: DprA Conveys Incoming ssDNA to RecA. Cell **130**:824–836.

81. **Bergé M**, **Mortier-Barrière I**, **Martin B**, **Claverys JP**. 2003. Transformation of Streptococcus pneumoniae relies on DprA- and RecA-dependent protection of incoming DNA single strands. Mol Microbiol **50**:527–536.

82. **Chen Z**, **Yang H**, **Pavletich NP**. 2008. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature **453**:489–484.

83. **Håvarstein LS**, **Coomaraswamy G**, **Morrison DA**. 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in Streptococcus pneumoniae. Proc Natl Acad Sci U S A **92**:11140–4.

84. **Claverys J**, **Prudhomme M**, **Martin B**. 2006. Induction of competence regulons as a general response to stress in gram-positive bacteria. Annu Rev Microbiol **60**:451–475.

85. **Sisco KL**, **Smith HO**. 1979. Sequence-specific DNA uptake in Haemophilus transformation. Proc Natl Acad Sci U S A **76**:972–976.

86. **Graves JF**, **Biswas GD**, **Sparling PF**. 1982. Sequence-specific DNA uptake in transformation of Neisseria gonorrhoeae. J Bacteriol.

87. **Dalia AAB**, **McDonough E**, **Camilli A**. 2014. Multiplex genome editing by natural transformation. Proc Natl Acad Sci U S A **111**:8937–42.

88. **Cooper TF**. 2007. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of Escherichia coli. PLoS Biol **5**:1899–1905.

89. **Haynes LL**, **Rothman-Denes LB**. 1985. N4 virion RNA polymerase sites of transcription initiation. Cell **41**:597–605.

90. **Esvelt KM**, **Wang HH**. 2013. Genome-scale engineering for systems and synthetic biology. Mol Syst Biol **9**:641.

91. **Pál C**, **Papp B**, **Pósfai G**. 2014. The dawn of evolutionary genome engineering. Nat Rev Genet **15**:504–12.

92. **Pósfai G**, **Umenhoffer K**, **Kolisnychenko V**, **Stahl B**, **Sharma SS**. 2008. Emergent Properties of Reduced-Genome *Escherichia coli* **1044**:1044–1047.

93. **Lajoie MJ**, **Rovner AJ**, **Goodman DB**, **Aerni H**, **Haimovich AD**, **Kuznetsov G**, **Mercer J a**, **Wang HH**, **Carr P a**, **Mosberg J a**, **Rohland N**, **Schultz PG**, **Jacobson JM**, **Rinehart J**, **Church GM**, **Isaacs FJ**. 2013. Genomically Recoded Organisms Expand Biological Functions. Science **342**:357–360.

94. **de Berardinis V**, **Durot M**, **Weissenbach J**, **Salanoubat M**. 2009. Acinetobacter baylyi ADP1 as a model for metabolic system biology. Curr Opin Microbiol **12**:568–576.

95. **Palmen R**, **Vosman B**, **Buijsman P**, **Breek CK**, **Hellingwerf KJ**. 1993. Physiological characterization of natural transformation in Acinetobacter calcoaceticus. J Gen Microbiol **139**:295–305.

96. **Simpson DJ**, **Dawson LF**, **Fry JC**, **Rogers HJ**, **Day MJ**. 2007. Influence of flanking homology and insert size on the transformation frequency of Acinetobacter baylyi BD413. Environ Biosafety Res **6**:55–69.

97. **Datsenko K a**, **Wanner BL**. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97**:6640–6645.

98. **Bacher JM**, **Metzgar D**, **De Crécy-Lagard V**. 2006. Rapid evolution of diminished transformability in Acinetobacter baylyi. J Bacteriol **188**:8534–8542.

99. **Umenhoffer K**, **Fehér T**, **Balikó G**, **Ayaydin F**, **Pósfai J**, **Blattner FR**, **Pósfai G**. 2010. Reduced evolvability of Escherichia coli MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. Microb Cell Fact **9**:38.

100. **Darmon E**, **Leach DRF**. 2014. Bacterial genome instability. Microbiol Mol Biol Rev **78**:1–39.

101. **de Berardinis V**, **Vallenet D**, **Castelli V**, **Besnard M**, **Pinet A**, **Cruaud C**, **Samair S**, **Lechaplais C**, **Gyapay G**, **Richez C**, **Durot M**, **Kreimeyer A**, **Le Fèvre F**, **Schächter V**, **Pezo V**, **Döring V**, **Scarpelli C**, **Médigue C**, **Cohen GN**, **Marlière P**, **Salanoubat M**, **Weissenbach J**. 2008. A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1. Mol Syst Biol **4**:174.

102. **Dong D**, **Yan A**, **Liu H**, **Zhang X**, **Xu Y**. 2006. Removal of humic substances from soil DNA using aluminium sulfate. J Microbiol Methods **66**:217–222.

103. **Baker GC**, **Smith JJ**, **Cowan DA**. 2003. Review and re-analysis of domain-specific 16S primers. J Microbiol Methods **55**:541–555.

104. **Murin CD**, **Segal K**, **Bryksin A**, **Matsumura I**. 2012. Expression vectors for Acinetobacter baylyi ADP1. Appl Environ Microbiol **78**:280–283.

105. **Tenaillon O**, **Taddei F**, **Radman M**, **Matic I**. 2001. Second-order selection in bacterial evolution: Selection acting on mutation and recombination rates in the course of adaptation. Res Microbiol **152**:11–16.

106. **Barbe V**, **Vallenet D**, **Fonknechten N**, **Kreimeyer A**, **Oztas S**, **Labarre L**, **Cruveiller S**, **Robert C**, **Duprat S**, **Wincker P**, **Ornston LN**, **Weissenbach J**, **Marlière P**, **Cohen GN**, **Médigue C**. 2004. Unique features revealed by the genome sequence of Acinetobacter sp. ADP1, a versatile and naturally transformation competent bacterium. Nucleic Acids Res **32**:5766–5779.

107. **Deatherage DE**, **Barrick JE**. 2014. Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol **1151**:165–188.

108. **Zerbino D**, **Birney E**. 2008. Velvet : Algorithms for De Novo Short Read Assembly Using De Bruijn Graphs. Most **18**:821–829.

109. **Altschul S**, **Madden T**, **Schaffer A**, **Zhang J**, **Zhang Z**, **Miller W**, **Dj L**. 1997. Gapped BLAST and PSI- BLAST: a new generation of protein database search programs. Nucleic acids Res **25**:3389–3402.

110. **Team RDC**. 2014. R: a language and environment for statistical computing. 3.1.0. The R Foundation for Statistical Computing, Vienna, Austria.

111. **Toren A**, **Navon-Venezia S**, **Ron EZ**, **Rosenberg E**. 2001. Emulsifying Activities of Purified Alasan Proteins from Acinetobacter radioresistens KA53. Appl Environ Microbiol **67**:1102–1106.

112. **Kibota TT**, **Lynch M**. 1996. Estimate of the genomic mutation rate deleterious to overall fitness in E. coli. Nature **381**:694–696.

113. **Barrick JE**, **Lenski RE**. 2013. Genome dynamics during experimental evolution. Nat Rev Genet **14**:827–39.

114. **Zolkiewski M**. 2006. A camel passes through the eye of a needle: Protein unfolding activity of Clp ATPases. Mol Microbiol **61**:1094–1100.

115. **Gottesman S**, **Clark WP**, **Maurizi MR**. 1990. The ATP-dependent Clp protease of Escherichia coli. Sequence of clpA and identification of a Clp-specific substrate. J Biol Chem **265**:7886–7893.

116. **Segura A**, **Bünz P V**, **D'Argenio D a**, **Ornston LN**. 1999. Genetic analysis of a chromosomal region containing *vanA* and *vanB*, genes required for conversion of

either ferulate or vanillate to protocatechuate in *Acinetobacter*. J Bacteriol **181**:3494–504.

117. **Nakar D**, **Gutnick DL**. 2001. Analysis of the wee gene cluster responsible for the biosynthesis of the polymeric bioemulsifier from the oil-degrading strain Acinetobacter Iwoffii RAG-1. Microbiology **147**:1937–1946.

118. **Vaneechoutte M**, **Tjernberg I**, **Baldi F**, **Pepi M**, **Fani R**, **Sullivan ER**, **Toorn J Van Der**, **Dijkshoorn L**. 1999. Oil-degrading Acinetobacter strain RAG-1 and strains described as "Acinetobacter venetianus sp. nov." belong to the same genomic species. Res Microbiol **150**:69–73.

119. **Kaplan N**, **Zosim Z**, **Rosenberg E**. 1987. Reconstitution of emulsifying activity of *Acinetobacter calcoaceticus* BD4 emulsan by using pure polysaccharide and protein. Appl Environ Microbiol **53**:440–446.

120. **Gerischer U**, **D'Argenio D a**, **Ornston LN**. 1996. IS*1236*, a newly discovered member of the IS*3* family, exhibits varied patterns of insertion into the *Acinetobacter calcoaceticus* chromosome. Microbiology **142**:1825–1831.

121. **Duval-Valentin G**, **Marty-Cointin B**, **Chandler M**. 2004. Requirement of IS911 replication before integration defines a new bacterial transposition pathway. EMBO J **23**:3897–3906.

122. **Cuff LE**, **Elliott KT**, **Seaton SC**, **Ishaq MK**, **Laniohan NS**, **Karls AC**, **Neidle EL**. 2012. Analysis of IS1236-mediated gene amplification events in Acinetobacter baylyi ADP1. J Bacteriol **194**:4395–4405.

123. **Jezequel N**, **Lagomarsino MC**, **Heslot F**, **Thomen P**. 2013. Long-term diversity and genome adaptation of Acinetobacter baylyi in a minimal-medium chemostat. Genome Biol Evol **5**:87–97.

124. **Welty DJ**, **Jones JM**, **Nakai H**. 1997. Communication of ClpXP protease hypersensitivity to bacteriophage Mu repressor isoforms. J Mol Biol **272**:31–41.

125. **Sekine Y**, **Izumi K**, **Mizuno T**, **Ohtsubo E**. 1997. Inhibition of transpositional recombination by OrfA and OrfB proteins encoded by insertion sequence IS*3*. Genes Cells **2**:547–557.

126. **Philippe N**, **Crozat E**, **Lenski RE**, **Schneider D**. 2007. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. BioEssays **29**:846–860.

127. **Conrad TM**, **Lewis NE**, **Palsson BO**. 2011. Microbial laboratory evolution in the era of genome-scale science. Mol Syst Biol **7**:509.

128. **Barrick JE**, **Yu DS**, **Yoon SH**, **Jeong H**, **Oh TK**, **Schneider D**, **Lenski RE**, **Kim JF**. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature **461**:1243–1247.

129. **Palmen R**, **Hellingwerf KJ**. 1995. *Acinetobacter calcoaceticus* liberates chromosomal DNA during induction of competence by cell lysis. Curr Microbiol **30**:7–10.

130. **Williams KP**. 2002. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucleic Acids Res **30**:866–875.

131. **Cooper VS**, **Schneider D**, **Blot M**, **Lenski RE**. 2001. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. J Bacteriol **183**:2834–2841.

132. **Renda BA**, **Hammerling MJ**, **Barrick JE**. 2014. Engineering reduced evolutionary potential for synthetic biology. Mol Biosyst **10**:1668–1678.

133. **Murin CD**, **Segal K**, **Bryksin A**, **Matsumura I**. 2011. Expression vectors for the engineering of genes and genomes in *Acinetobacter baylyi* ADP1. Appl Environ Microbiol **78**:280–283.

134. **Casjens S**. 2003. Prophages and bacterial genomics: What have we learned so far? Mol Microbiol **49**:277–300.

135. **Canchaya C**, **Fournous G**, **Brüssow H**. 2004. The impact of prophages on bacterial chromosomes. Mol Microbiol **53**:9–18.

136. **Wang X**, **Kim Y**, **Ma Q**, **Hong SH**, **Pokusaeva K**, **Sturino JM**, **Wood TK**. 2010. Cryptic prophages help bacteria cope with adverse environments. Nat Commun **1**:147.

137. **Bossi L**, **Fuentes J a**, **Mora G**, **Figueroa-bossi N**. 2003. Prophage Contribution to Bacterial Population Dynamics Prophage Contribution to Bacterial Population Dynamics. J Bacteriol **185**:6467–6471.

138. **Lang AS**, **Zhaxybayeva O**, **Beatty JT**. 2012. Gene transfer agents: phage-like elements of genetic exchange. Nat Rev Microbiol **10**:472–482.

139. **Waldor MK**, **Mekalanos JJ**. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science **272**:1910–1914.

140. **Lorenz MG**, **Wackernagel W**. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev **58**:563–602.

141. **Averhoff B**, **Friedrich A**. 2003. Type IV pili-related natural transformation systems: DNA transport in mesophilic and thermophilic bacteria. Arch Microbiol **180**:385–393.

142. **Dubnau D**. 1991. The regulation of genetic competence in Bacillus subtilis. Mol Microbiol **5**:11–18.

143. **Palmen R**, **Vosman B**, **Buijsman P**, **Breek CK**, **Hellingwerf KJ**. 1993. Physiological characterization of natural transformation in *Acinetobacter calcoaceticus*. J Gen Microbiol **139**:295–305.

144. **Renda BA**, **Dasgupta A**, **Leon D**, **Barrick JE**. 2015. Genome Instability Mediates the Loss of Key Traits by Acinetobacter baylyi ADP1 during Laboratory Evolution. J Bacteriol **197**:872–881.

145. **Russel M**, **Model P**. 2006. Filamentous Phage, p. 146–160. *In* The Bacteriophages Second Edition.

146. **Redfield RJ**. 2001. Do bacteria have sex? Nat Rev Genet **2**:634–639.

147. **Baltrus D a.**, **Guillemin K**, **Phillips PC**. 2008. Natural transformation increases the rate of adaptation in the human pathogen Helicobacter pylori. Evolution (N Y) **62**:39–49.

148. **Sambrook J**, **Russell DW**. 2001. Preparation of Single-stranded Bacteriophage M13 DNA, p. 3.26. *In* Molecular Cloning: A Laboratory Manual Third Edition.

149. **Barrick JE**, **Colburn G**, **Deatherage DE**, **Traverse CC**, **Strand MD**, **Borges JJ**, **Knoester DB**, **Reba A**, **Meyer AG**. 2014. Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. BMC Genomics **15**:1039.

150. **Casjens S**, **Winn-Stapley DA**, **Gilcrease EB**, **Morona R**, **Kuhlewein C**, **Chua JEH**, **Manning PA**, **Inwood W**, **Clark AJ**. 2004. The chromosome of Shigella flexneri bacteriophage Sf6: Complete nucleotide sequence, genetic mosaicism, and DNA packaging. J Mol Biol **339**:379–394.

151. **Kearse M**, **Moir R**, **Wilson A**, **Stones-Havas S**, **Cheung M**, **Sturrock S**, **Buxton S**, **Cooper A**, **Markowitz S**, **Duran C**, **Thierer T**, **Ashton B**, **Meintjes P**, **Drummond A**. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28**:1647–1649.

152. **Davis BM**, **Moyer KE**, **Fidelma Boyd E**, **Waldor MK**. 2000. CTX prophages in classical biotype Vibrio cholerae: Functional phage genes but dysfunctional phage genomes. J Bacteriol **182**:6992–6998.

153. **Waldor MK**, **Rubin EJ**, **Pearson GD**, **Kimsey H**, **Mekalanos JJ**. 1997. Regulation, replication, and integration functions of the Vibrio cholerae CTXphi are encoded by region RS2. Mol Microbiol **24**:917–926.

154. **Davis BM**, **Kimsey HH**, **Kane A V.**, **Waldor MK**. 2002. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO J **21**:4240–4249.

155. **Fasano A**, **Baudry B**, **Pumplin DW**, **Wasserman SS**, **Tall BD**, **Ketley JM**, **Kaper JB**. 1991. *Vibrio cholerae* produces a second enterotoxin, which affects intestinal tight junctions. Proc Natl Acad Sci U S A **88**:5242–5246.

156. **Koonin E V.** 1992. The second cholera toxin, Zot, and its plasmid-encoded and phage-encoded homologues constitute a group of putative ATPases with an altered purine NTP-binding motif. FEBS Lett **312**:3–6.

157. **Trucksis M**, **Galen JE**, **Michalski J**, **Fasano A**, **Kaper JB**. 1993. Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. Proc Natl Acad Sci U S A **90**:5267–5271.

158. **Pearson GD**, **Woods a**, **Chiang SL**, **Mekalanos JJ**. 1993. CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. Proc Natl Acad Sci U S A **90**:3750–4.

159. **Meyer JR**, **Dobias DT**, **Weitz JS**, **Barrick JE**, **Quick RT**, **Lenski RE**. 2012. Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science (80-) **335**:428–432.

160. **Paterson S**, **Vogwill T**, **Buckling A**, **Benmayor R**, **Spiers AJ**, **Thomson NR**, **Quail M**, **Smith F**, **Walker D**, **Libberton B**, **Fenton A**, **Hall N**, **Brockhurst MA**. 2010. Antagonistic coevolution accelerates molecular evolution. Nature **464**:275–278.

161. **Perry EB**, **Barrick JE**, **Bohannan BJM**. 2015. The Molecular and Genetic Basis of Repeatable Coevolution between Escherichia coli and Bacteriophage T3 in a Laboratory Microcosm. PLoS One **10**:e0130639.

162. **Waldor MK**, **Rubin EJ**, **Pearson GD**, **Kimsey H**, **Mekalanos JJ**. 1997. Regulation, replication, and integration functions of the *Vibrio cholerae* CTXφ are encoded by region RS2. Mol Microbiol **24**:917–26.

163. **Kimsey HH**, **Waldor MK**. 2004. The CTX?? Repressor RstR Binds DNA Cooperatively to Form Tetrameric Repressor-Operator Complexes. J Biol Chem **279**:2640–2647.

164. **Bailone A**, **Devoret R**, **Riffell SK**, **Gutzwillerj KJ**. 1987. Isolation of Ultravirulent mutants of Phage Lambda. Virology **84**:547–550.

165. **Porstendörfer D**, **Drotschmann U**, **Averhoff B**. 1997. A novel competence gene, *comP*, is essential for natural transformation of *Acinetobacter* sp. strain BD413. Appl Environ Microbiol **63**:4150–7.

166. **Gohl O**, **Friedrich A**, **Hoppert M**, **Averhoff B**. 2006. The thin pili of Acinetobacter sp. strain BD413 mediate adhesion to biotic and abiotic surfaces. Appl Environ Microbiol **72**:1394–401.

167. **Henrichsen J**, **Blom J**. 1975. Correlation between twitching motility and possession of polar fimbriae in Acinetobacter calcoaceticus. Acta Pathol Microbiol Scand B **83**:103–115.

168. **Sahu SN**, **Acharya S**, **Tuminaro H**, **Patel I**, **Dudley K**, **LeClerc JE**, **Cebula T a.**, **Mukhopadhyay S**. 2003. The bacterial adaptive response gene, barA, encodes a novel conserved histidine kinase regulatory switch for adaptation and modulation of metabolism in Escherichia coli. Mol Cell Biochem **253**:167–177.

169. **Vakulskas CA**, **Potts AH**, **Babitzke P**, **Ahmer BM**, **Romeo T**. 2015. Regulation of bacterial virulence by Csr (Rsm) systems. Microbiol Mol Biol Rev **79**:193–224.

170. **Chavez RG**, **Alvarez AF**, **Romeo T**, **Georgellis D**. 2010. The physiological stimulus for the BarA sensor kinase. J Bacteriol **192**:2009–2012.

171. **Pires DP**, **Oliveira H**, **Melo LDR**, **Sillankorva S**, **Azeredo J**. 2016. Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl Microbiol Biotechnol 2141–2151.

172. **Labrie SJ**, **Samson JE**, **Moineau S**. 2010. Bacteriophage resistance mechanisms. Nat Rev Microbiol **8**:317–327.

173. **Harding CM**, **Tracy EN**, **Carruthers MD**, **Rather PN**, **Actis LA**, **Munson RS**. 2013. Acinetobacter baumannii strain M2 produces type IV Pili which play a role in natural transformation and twitching motility but not surface-associated motility. MBio **4**:1–10.

174. **Wilharm G**, **Piesker J**, **Laue M**, **Skiebe E**. 2013. DNA Uptake by the nosocomial pathogen acinetobacter baumannii occurs during movement along wet surfaces. J Bacteriol **195**:4146–4153.

175. **Touchon M**, **Cury J**, **Yoon E-J**, **Krizova L**, **Cerqueira GC**, **Murphy C**, **Feldgarden M**, **Wortman J**, **Clermont D**, **Lambert T**, **Grillot-Courvalin C**, **Nemec A**, **Courvalin P**, **Rocha EPC**. 2014. The genomic diversification of the whole *Acinetobacter* genus: origins, mechanisms, and consequences. Genome Biol Evol **6**:2866–82.

176. **Croucher NJ**, **Mostowy R**, **Wymant C**, **Turner P**, **Bentley SD**, **Fraser C**. 2016. Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic Conflict. PLOS Biol **14**:e1002394.

177. **Yang H**, **Liang L**, **Lin S**, **Jia S**. 2010. Isolation and characterization of a virulent bacteriophage AB1 of Acinetobacter baumannii. BMC Microbiol **10**:131.

178. **Herman NJ**, **Juni E**. 1974. Isolation and characterization of a generalized transducing bacteriophage for Acinetobacter. J Virol **13**:46–52.

179. **Lin NT**, **Chiou PY**, **Chang KC**, **Chen LK**, **Lai MJ**. 2010. Isolation and characterization of AB2: A novel bacteriophage of *Acinetobacter baumannii*. Res Microbiol **161**:308–314.

180. **Klovins J**, **Overbeek GP**, **van den Worm SHE**, **Ackermann HW**, **van Duin J**. 2002. Nucleotide sequence of a ssRNA phage from *Acinetobacter*: Kinship to coliphages. J Gen Virol **83**:1523–1533.

181. **Palchoudhury SR**, **Iyer VN**. 1969. Loss of an episomal fertility factor following the multiplication of coliphage M13. MGG Mol Gen Genet **105**:131–139.

182. **Biswas GD**, **Burnstein KL**, **Sparling PF**. 1986. Linearization of donor DNA during plasmid transformation in Neisseria gonorrhoeae. J Bacteriol **168**:756–761.

183. **Shen P**, **Huang H V.** 1986. Homologous recombination in Escherichia coli: dependence on substrate length and homology. Genetics **112**:441–457.

184. **Gibson DG**, **Young L**, **Chuang R-Y**, **Venter JC**, **Hutchison C a**, **Smith HO**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods **6**:343–345.

185. **Tucker AT**, **Nowicki EM**, **Boll JM**, **Knauf GA**, **Burdis NC**, **Stephen Trent M**, **Davies BW**. 2014. Defining gene-phenotype relationships in acinetobacter baumannii through one-step chromosomal gene inactivation. MBio **5**:1–9.

186. **Register JC**, **Griffith J**. 1985. The direction of RecA protein assembly onto single strand DNA is the same as the direction of strand assimilation during strand exchange. J Biol Chem **260**:12308–12312.

187. **Letters T**. 1985. Nucleoside phosphorothioates. J Am Chem Soc 4292–4294.

188. **Gallie DR**, **Sleat DE**, **Watts JW**, **Turner PC**, **Wilson TMA**. 1988. Inhibition of deoxyribonucleases by phosphorothioate groups in oligodeoxyribonucleotides. Nucleic acids Res **16**:883–893.

189. **Jasin M**, **Schimmel P**. 1984. Deletion of an essential gene in Escherichia coli by site-specific recombination with linear DNA fragments. J Bacteriol **159**:783–786.

190. **Claverys JP**, **Martin B**, **Polard P**. 2009. The genetic transformation machinery: Composition, localization, and mechanism. FEMS Microbiol Rev **33**:643–656.

191. **Mirouze N**, **Bergé MA**, **Soulet A-L**, **Mortier-Barrière I**, **Quentin Y**, **Fichant G**, **Granadel C**, **Noirot-Gros M-F**, **Noirot P**, **Polard P**, **Martin B**, **Claverys J-P**. 2013. Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. Proc Natl Acad Sci U S A **110**:E1035–44.

192. **de Vries J**, **Wackernagel W**. 2002. Integration of foreign DNA during natural transformation of Acinetobacter sp. by homology-facilitated illegitimate recombination. Proc Natl Acad Sci U S A **99**:2094–2099.

193. **van Opijnen T**, **Camilli A**. 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol **11**:435–42.

194. **Kulasekara HD**, **Ventre I**, **Kulasekara BR**, **Lazdunski A**, **Filloux A**, **Lory S**. 2005. A novel two-component system controls the expression of Pseudomonas aeruginosa fimbrial cup genes. Mol Microbiol **55**:368–380.

195. **Demarre G**, **Guérout AM**, **Matsumoto-Mashimo C**, **Rowe-Magnus DA**, **Marlière P**, **Mazel D**. 2005. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPα) conjugative machineries and their cognate Escherichia coli host strains. Res Microbiol **156**:245–255.

196. **Langmead B**, **Salzberg SL**. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods **9**:357–359.

197. **Love MI**, **Huber W**, **Anders S**. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**:550.

198. **Turner KH**, **Wessel AK**, **Palmer GC**, **Murray JL**, **Whiteley M**. 2015. Essential genome of Pseudomonas aeruginosa in cystic fibrosis sputum. Proc Natl Acad Sci U S A **112**:4110–5.

199. **Lee ME**, **DeLoache WC**, **Cervantes B**, **Dueber JE**. 2015. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. ACS Synth Biol **4**:975–986.

200. **Chao MC**, **Abel S**, **Davis BM**, **Waldor MK**. 2016. The design and analysis of transposon insertion sequencing experiments. Nat Rev Microbiol **14**:119–128.

201. **Mattick JS**. 2002. Type IV pili and twitching motility. Annu Rev Microbiol **56**:289–314.

202. **Hobbs M**, **Collie ES**, **Free PD**, **Livingston SP**, **Mattick JS**. 1993. PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in Pseudomonas aeruginosa. Mol Microbiol **7**:669–82.

203. **Ishimoto KS**, **Lory S**. 1992. Identification of pilR, which encodes a transcriptional activator of the Pseudomonas aeruginosa pilin gene. J Bacteriol **174**:3514–3521.

204. **McNeil MB**, **Clulow JS**, **Wilf NM**, **Salmond GPC**, **Fineran PC**. 2012. SdhE is a conserved protein required for flavinylation of succinate dehydrogenase in bacteria. J Biol Chem **287**:18418–18428.

205. **McNeil MB**, **Iglesias-Cans MC**, **Clulow JS**, **Fineran PC**. 2013. YgfX (CptA) is a multimeric membrane protein that interacts with the succinate dehydrogenase assembly factor sdhe (YgfY). Microbiol (United Kingdom) **159**:1352–1365.

206. **Masuda H**, **Tan Q**, **Awano N**, **Yamaguchi Y**, **Inouye M**. 2012. A novel membrane-bound toxin for cell division, CptA (YgfX), inhibits polymerization of cytoskeleton proteins, FtsZ and MreB, in Escherichia coli. FEMS Microbiol Lett **328**:174–181.

207. **Jin S**, **Ishimoto KS**, **Lory S**. 1994. PilR, a transcriptional regulator of piliation in Pseudomonas aeruginosa, binds to a cis-acting sequence upstream of the pilin gene promoter. Mol Microbiol **14**:1049–1057.