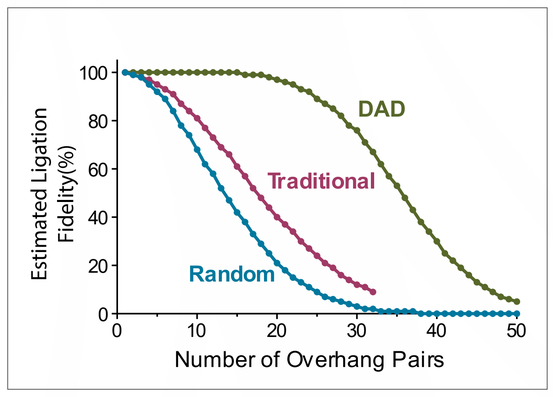
Application of High Complexity Golden Gate Assembly to Rapid Engineering of Bacteriophage Genomes

Andrew Sikkema, Katharina Bilotti, S. Kasra Tabatabaei, Sean Lund, Yan-Jiun Lee, Vladimir Potapov, **Gregory J. S. Lohman** New England Biolabs, Inc.

Data-optimized Assembly Design – Application of Ligase Fidelity to Golden Gate Assembly



Data-optimized Assembly Design (DAD) applies comprehensive measurements of ligation fidelity to flexibly select high-accuracy sets of fusion sites for Golden Gate Assembly (GGA). DAD replaces classic GGA rules of thumb for junction selection or the need to rely on pre-validated sets, permitting highly accurate and efficient assembly of dozens of parts in a single Golden Gate reaction while minimizing possible mis-assembly events. We have developed a set of tools to apply DAD to the design of complex assemblies of up to 36 parts and 50kB final size in a single reaction. When used in combination with optimized reagents and protocols, high yields of complete, sequence accurate constructs can be achieved, permitting direct use through standard transformation and rescue protocols without purification or other post-assembly processing steps. The parts are small enough to be conveniently produced via PCR or DNA synthesis and are rationally designed to produce fragments that, with few exceptions, are readily propagated and manipulated in traditional *E. coli* plasmid systems. Small, easily propagated parts support easy viral mutagenesis via classic molecular biology methods as well as inexpensive whole gene swaps by substituting new parts. We have applied these principles to develop systems for the modular, one-step assembly of bacteriophage genomes, including several systems targeting human pathogens with potential therapeutic applications. These synthetic chassis will be used in programs of high throughput study of genotype/phenotype host range connections.

Ligation fidelity measurements can be used to predict high fidelity and low fidelity fusion site sets with confidence.

High-fidelity DAD enables:

- **More colonies/plaques**
- **Low error rate/sequence accuracy**
- **Minimal screening**
- **Many fragments in one reaction**
- **Direct transformation without purification or passage through a secondary organism**

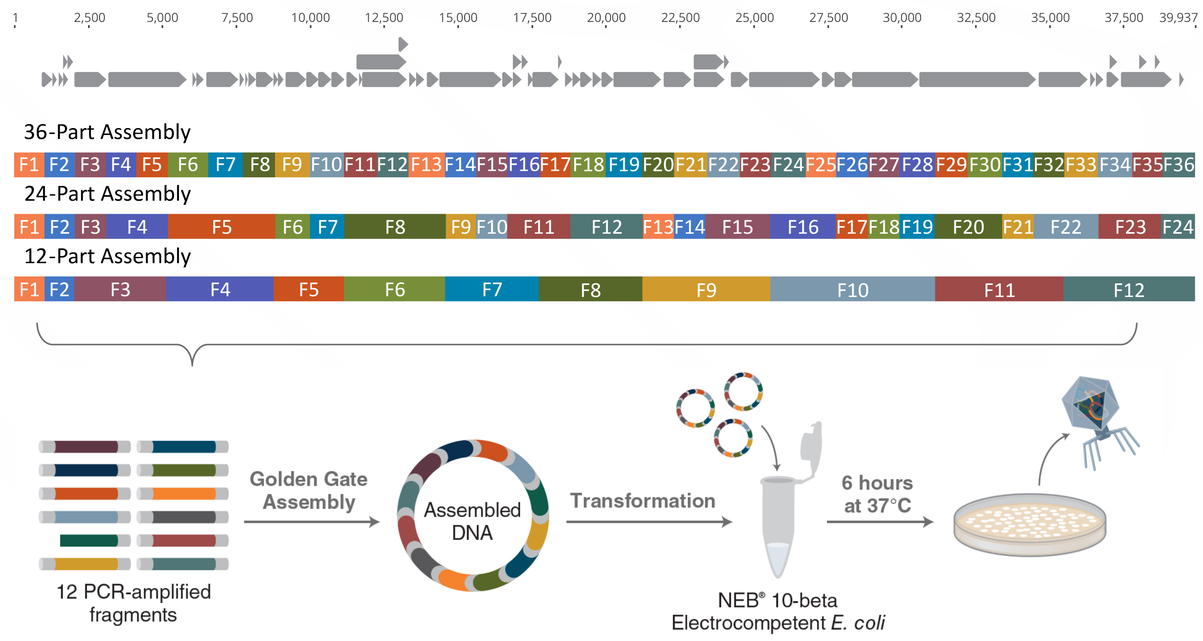
Estimated Fidelity 100%

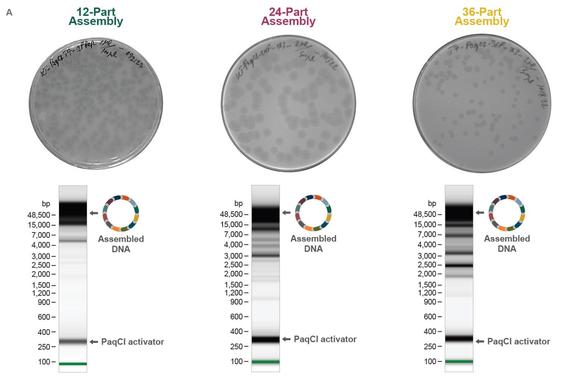
	AAGG	CCTT	GCTA	TAGC	CTCC	GGAG	ACGA	TCGT
AAGG		■						
CCTT	■							
GCTA				■				
TAGC			■					
CTCC						■		
GGAG					■			
ACGA								■
TCGT							■	

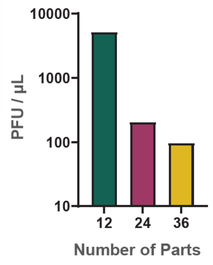
Estimated Fidelity 37%

	GTCA	TGAC	CGGC	GCCG	CGTC	GACG	CGAC	GTCG
GTCA		■					■	
TGAC	■							■
CGGC				■		■		■
GCCG			■		■		■	
CGTC				■		■		■
GACG			■		■		■	
CGAC	■			■		■		■
GTCG		■	■		■		■	

T7 Bacteriophage genome assembly from up to 36 parts







- **Parts to plaques in <2 days**
- **100% of selected plaques contain desired domesticating mutations**
- **Error rate limited by PCR (~1/3 plaques contain a single random SNP)**
- **Doubling number of parts decreases PFU/µL ~10x**
- **DNA quality critical – a single PCR amplicon with primer dimer contamination drops PFU/µL ~10x**

Keys to Assembly Success:

- **High fidelity set of fusion sites**
- **Clean, sequence verified plasmids**
- **PCR amplicons free of primer dimers**
- **Equimolar mixtures of parts**
- **Optimized reagents and protocols**

NEBridge® Golden Gate Assembly Tool
https://goldengate.neb.com

NEBridge® Ligase Fidelity Tools for Golden Gate Assembly
https://ligasefidelity.neb.com

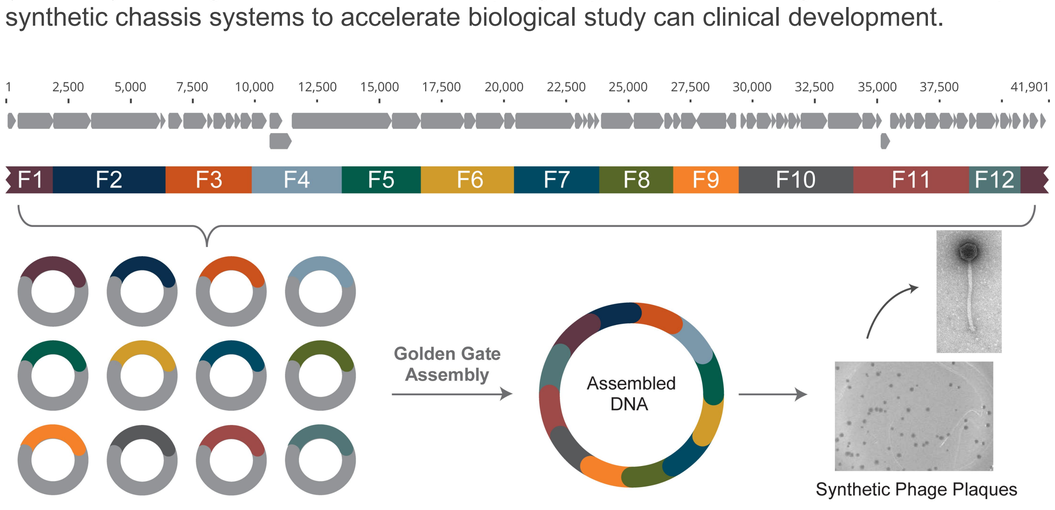
High-Complexity One-Pot Golden Gate Assembly
Andrew P. Sikkema, S. Kasra Tabatabaei, Yan-Jiun Lee, Sean Lund, Gregory J. S. Lohman
https://doi.org/10.1002/cpz1.882

Lohman Lab @NEB

High GC *Mycobacterium smegmatis* Phage BPs

ANSA BIOTECHNOLOGIES
Jeremiah Hanes
Nicole Chew
Tan Li
Yufeng Qian
University of Pittsburgh
Graham Hatfull
Ching-Chung Ko

Mycobacteria are Gram-positive organisms with GC rich (60-70%) genomes. Numerous species, including *M. tuberculosis, M. abscessus and M. avium* are important human pathogens, while the BL1 sp. *M. smegmatis* is often studied in a laboratory context. Over 12,000 mycobacterial phages have been sequenced, thanks largely to the SEA-PHAGES program which has organized students and researchers around the world in isolating and sequencing these phages. This database offers unmatched diversity for potential identification and engineering these phages as potential therapeutic agents. With their large genome size (42-70kB) and high GC contents, these phages present a challenge to the development of synthetic chassis systems to accelerate biological study can clinical development.

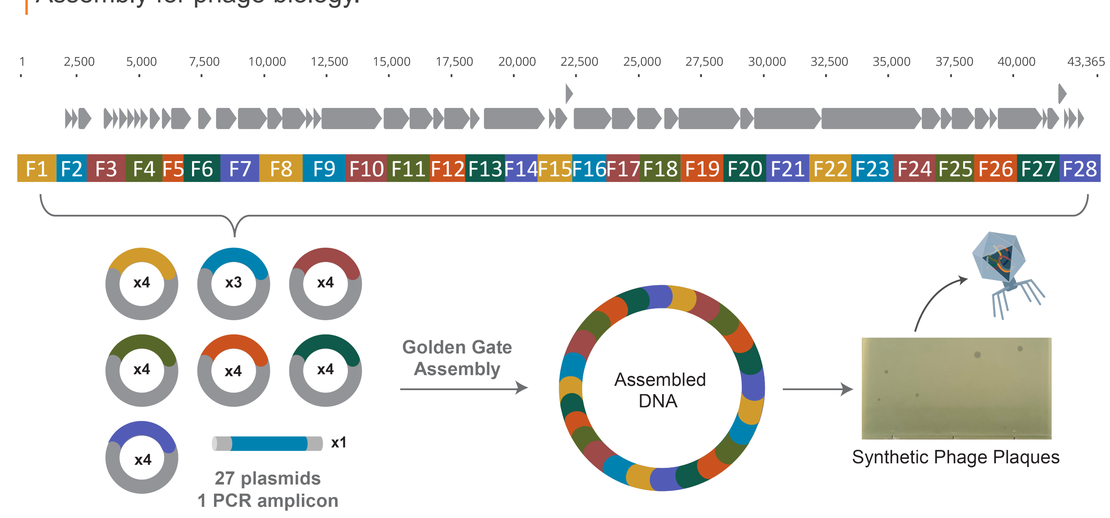


- **~67% GC genome with fragments ranging from 64-68.5%. Fragments could not be synthesized by current generation DNA vendors.**
- **Ansa provided all 12 fragments (2.1-4.7 kB) in 6 weeks from order, sequence verified in a GGA compatible holding vector.**
- **Assembly successful on first attempt; reboot likewise successful on first transformation of the assembly mixture.**
- **12/12 picked plaques were 100% sequence perfect with all intended mutations and no SNPs or indels.**
- **System represents a chassis useful for rapid investigation of genotype/phenotype investigation and a model for development of broad host range strains for phage therapeutics.**

Pseudomonas aeruginosa Phage ϕKMV

FELIX BIOTECHNOLOGY
Robert McBride
Max Hopkins
Hemaa Selvakumar
Yale University School of Medicine
Paul E. Turner
Matthew Davidson
Kaitlyn E. Kortright

Pseudomonas aeruginosa is an opportunistic human pathogen associated with hospital acquired infection and infection of immune compromised patients. Clinical strains of *P. aeruginosa* are commonly multi-drug resistant making *P. aeruginosa* a desirable target for phage-based therapies. However, the development of phages for clinical applications will require easy to manipulate systems for high throughput screening of variants for advantageous qualities like expand host range and pathogenicity. The phage ϕKMV is a T7-like bacteriophage targeting *Pseudomonas* that we selected as a model system to study the use of Golden Gate Assembly for phage biology.



- **Targets an ESKAPE pathogen generally associated with both nosocomial infections and infection of immunocompromised patients**
- **Fully synthetic assembly system with fragments <1800 bp**
- **Assembly product is directly transformed into *P. aeruginosa***
- **Point mutations generated via Site-Directed Mutagenesis of fragments**
- **20/20 plaques screened contained all desired mutations**
- **Only one plaque contains a single SNP in the PCR amplicon region.**
- **Phenotype/Genotype analysis of host range changes underway.**
- **System represents a chassis for the high throughput development of phage therapeutics against this important pathogen.**