Team

Album Shen  
BME '14

George Chen  
BME '15

Robert Eisinger  
Computer Science '13

Emily Scher  
Computer Science '15

Steffi Liu  
BME '13

Dr. Jef Boeke

Dr. Patrick Yizhi Cai

Dr. Joel S. Bader

Dr. Giovanni Stracquadanio
Do you speak DNA?

... It's not so easy.
Understanding DNA

By Dr. Seuss

One fish
two fish
red fish
blue fish
BioBricks

Year

Total Number of Biobricks

2003 2004 2005 2006 2007 2008 2009 2010 2011

Introduction AutoPlasmid Architecture AutoDesign Demo Case Study Future
Misannotations

Part: BBa_I732406

Promoter Family Member (U037011+D062NUL)
Promoter family member, For more details, please refer to BBa_I732400.

Sequence and Features

Assembly Compatibility: 10 12 21 23 25
Misannotations

Part: BBa_K122008

BleoR
Supf tRNA transcript under constitutive repressible transcription (replace with tyrosine)

Sequence and Features

Assembly Compatibility: 10 12 21 23 25

Parameters

Reviews

Categories

Introduction AutoPlasmid Architecture AutoDesign Demo Case Study Future
Widespread problem
Widespread problem

(Dramatization)
An unsolved problem: Imperfect matching

- Mutations are much more common in biology than they are in English
- Current software tools do not address this problem!
The Autogene Approach

- Speed, Efficiency & Correctness
- Features Database
- Automatic Annotation
- Plasmid Viewing

Introduction  AutoPlasmi d  Architecture  AutoDesign  Demo  Case Study  Future
Workflow

What’s in my DNA?

...AGCAACTCGGTCGCCGCATAAG
GCCAG...ACGGACAGAGACCCAGAG
AGACAGACCACCATTTCATTATTTTA
TAGGTTAATGTGCTGAAAATTAAACC
GATACCTTTGGGATATAGGTCTCAG
TCGCTGTGAGCTCTAGGTGTCGCGC
AAAACCCGTATTGAGCCGCGGCAAG
AGCAA...CTCGGAAAAGTGTCTCC
ATATTC...
Workflow

What’s in my DNA?

...AGCAACTCGGTCGCCGCATAAG
CCCAG...ACGGACAAGACACAGAG
AGACAGAGCCCGTTTCAATTATTATA
TAGGTTAATCTCAGAATTTAACC
GATACCTTCCGGGATATACCTCTCGA
TGGCTGTGAGCTCTAGGTGTCGCGC
AAAACCCGTATTGACGCGGGCAAG
ACCAA...CTCGGAAAAGGTGTCTCC
ATATTC...

Perfect Matches

Imperfect Matches
Workflow

What’s in my DNA?

...AGCAACTCGGTCGCCGCATAAG
GCCAG...ACGGACAAGAGCCACAG
AGACAGACCCAATTTTCTATTATTTTA
TAGGTTAATGTCATGAATTATAACC
GATACCTTGGGATATAGGTCTCGA
TCGGGTGTGAGCTACTGGGTGTCGCGC
AAAACCCGTATTGACGCCGGGCAAG
AGCAA...CTCGGAAAAGTGTCTCC
ATATTC...
Workflow

What’s in my DNA?

...AGCAACTCGGTGCAGCGCATATAG GCCAG...ACGGACAAGAGAGACAGAG AGACAGACCCCATTTCAATATTATTTTAGGTTAATGTCAATGAATTTAACC GATACTTGGAGATATAGGTCTCGA TCCCTGTGAGCTCTAGGTGTCGCGC AAAACCCGTATTGACGCCCCGGGCAAG AGCAA...CTCGGAAAAAGTGCTCTCC ATATTC...

| plasmid   | 1616 GCGTGACCTTTGAAAGCGCCACGCTTCCCAGGGAGAAAGGCCGACA | 1665 |
| feature   | 401 GCGTGACGATTGAAAGCGCCACGCTTCCCAGGGAGAAAGGCCGACA | 450  |
Software Architecture
What’s Under the Hood
Procedure

DNA

Feature
1
2
3
...

Match
✔
✖
✔

Database

Get Features

Algorithm

Introduction  AutoPlasmid  Architecture  AutoDesign  Demo  Case Study  Future
Difficult Procedure

Database

DNA

Get Features

Algo

plasmid  1616 GCGTGCTTTTGAAAGCGCCACGCTTTCCGAAGGAGAAAGGCGGACA  1665
feature  401 GCGTACGCAATTGAAAGAGCCACGCTTTCCGAAGGAGAAAGGCGGACA  450

Introduction  AutoPlasmid  Architecture  AutoDesign  Demo  Case Study  Future
Cloud Computing

Google

Database

Dropbox

amazon

facebook

Introduction  AutoPlasmi d  Architecture  AutoDesign  Demo  Case Study  Future
Architecture
Database

Virulence Factors Database
Architecture

[Diagram showing the relationship between a database, an algorithm, and a computer interface.]
Alignment Algorithm

- Smith-Waterman
- Exact algorithm
Architecture
Cloud Infrastructure
All You Need Is... Cloud Infrastructure
Cloud Performance

- PUC18: 2,680 bp
- Search against 18,000 features
Cloud Performance

PUC18: 2,680 bp
Search against 18,000 features

Improvement: 34x

Time (sec.)

Laptop

Subjobs

0 200 400 600 800 1000 1200 1400 2000 2340

39 min

1 2 5 10 15 20 25 30

68 sec
AutoDesign

From Words to Sentences
Demo

Autogene in Action
Case Study
The Autogene Solution
Pathogen Detection

• PROBLEM:
  There may be harmful BioBricks in the Registry

Solution 1

Good News: No Pathogens found; further analysis needed.

Solution 2
Future of Autogene

Auto Design

New BioBrick

BioBricks

New BioBrick

New BioBrick
Software Compatibility

• SBOL
• FASTA
• GenBank
• ApE
• gff3
• embl
Software Availability

Download Source Online:

Visit us at Poster SW#2 to receive a USB pre-loaded with Autogene!

(Limited supply)
Acknowledgements

- Carlos Olguin
- Florencio Mazzoldi
- Francesco Iorio
- Lawson Fulton
- Kimberly Whinna
- Glenn Katz
- Laura London
- Brian Pene
- Jeff Clayton
Future Features

- Codon optimization
- Primer design
- Protein visualization
- Pathogen analysis integration
- More file formats: gff3, embl
- GUI enhancements
- Common features bank
- Sequence indexing