jbei
Joint BioEnergy Institute

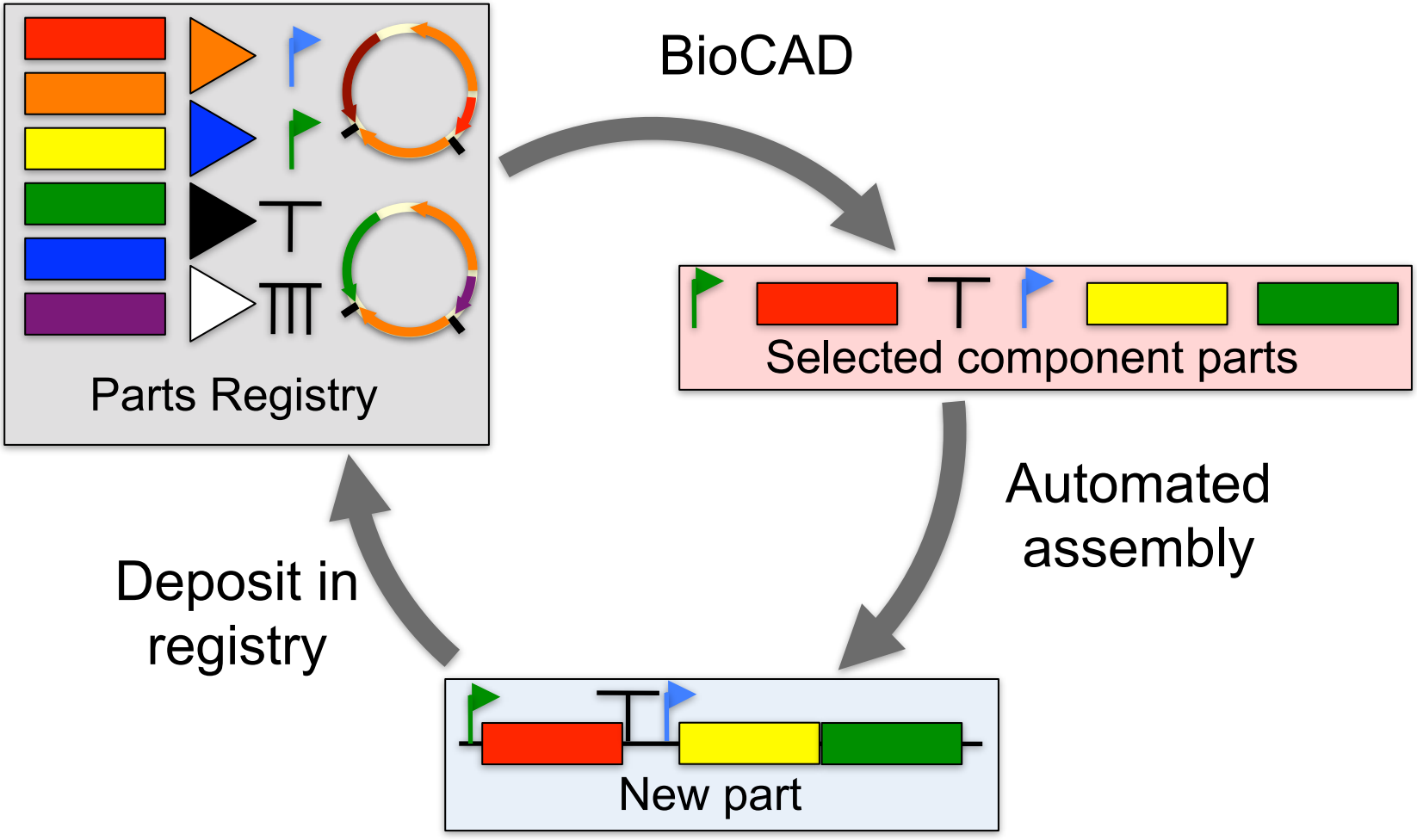
Parts registries, BioCAD, and DNA-assembly design automation

Nathan J. Hillson
njhillson@lbl.gov

NASA Ames Synthetic Biology Workshop
October 30, 2010

U.S. DEPARTMENT OF ENERGY Office of Science

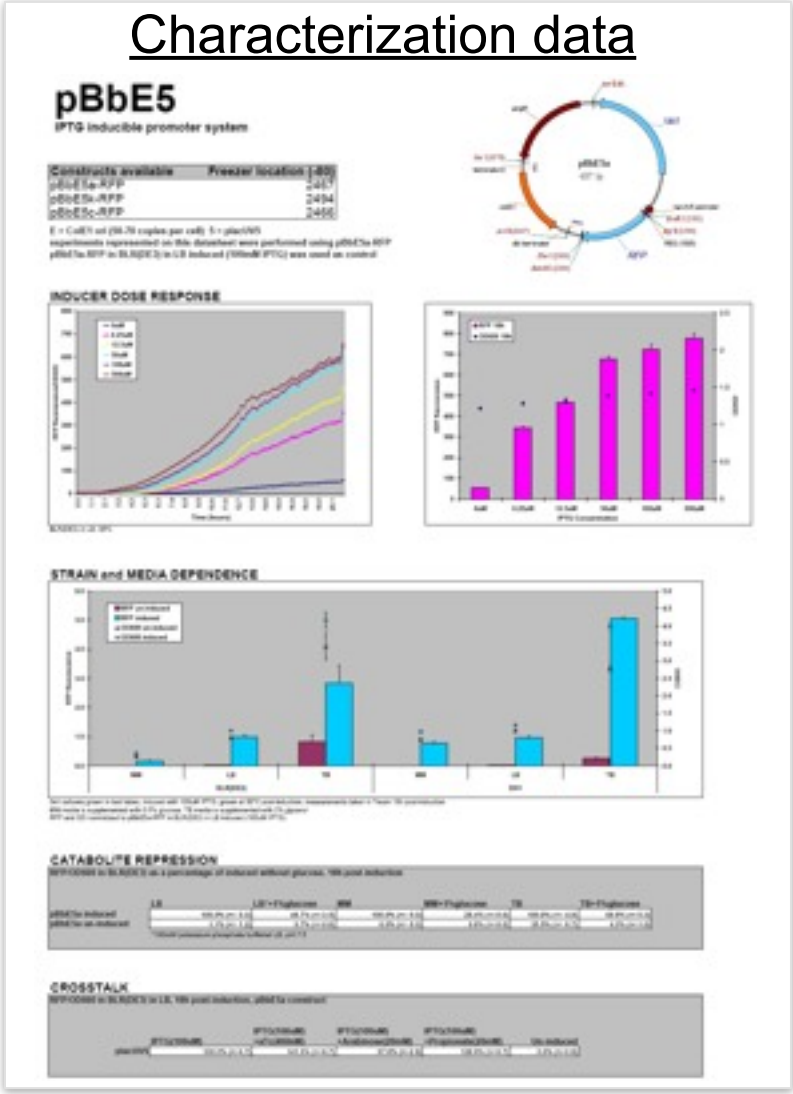
Synthetic Biology – an integrated process



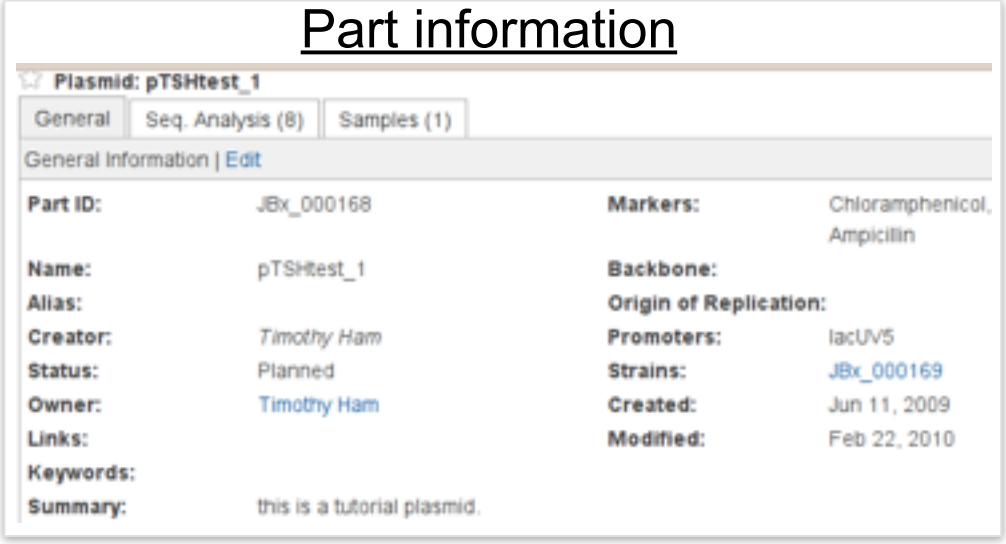
JBEI Registry of Biological Parts

- Advanced search features
- Machine-readable characterization data access in development
- Facilitates access to parts distributed across multiple labs
- Tracks intellectual property and funding sources

Characterization data



Part information



Plasmid: pTSHtest_1

General | Seq. Analysis (8) | Samples (1)

General Information | Edit

Part ID:	JBx_000168	Markers:	Chloramphenicol, Ampicillin
Name:	pTSHtest_1	Backbone:	
Alias:		Origin of Replication:	
Creator:	*Timothy Ham*	Promoters:	lacUV5
Status:	Planned	Strains:	JBx_000169
Owner:	Timothy Ham	Created:	Jun 11, 2009
Links:		Modified:	Feb 22, 2010
Keywords:			
Summary:	this is a tutorial plasmid.		

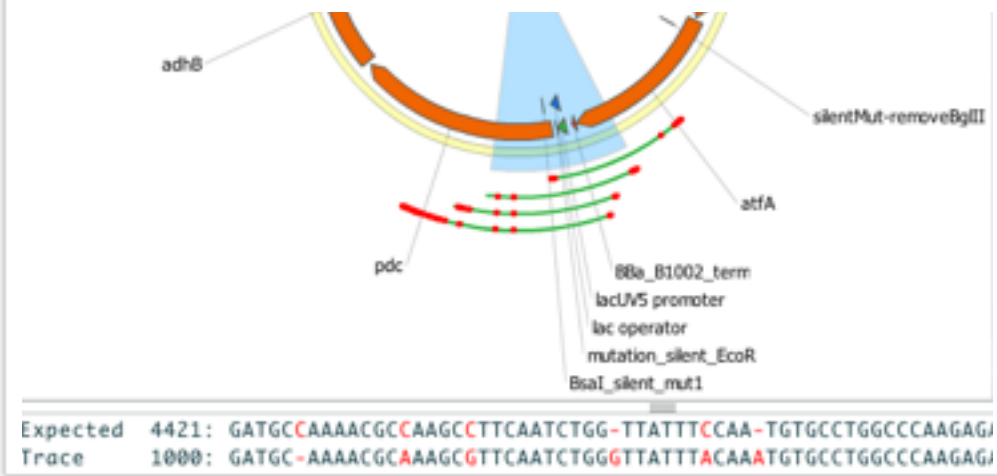
Timothy Ham, Zinovii Dmytriv

U.S. DEPARTMENT OF ENERGY Office of Science

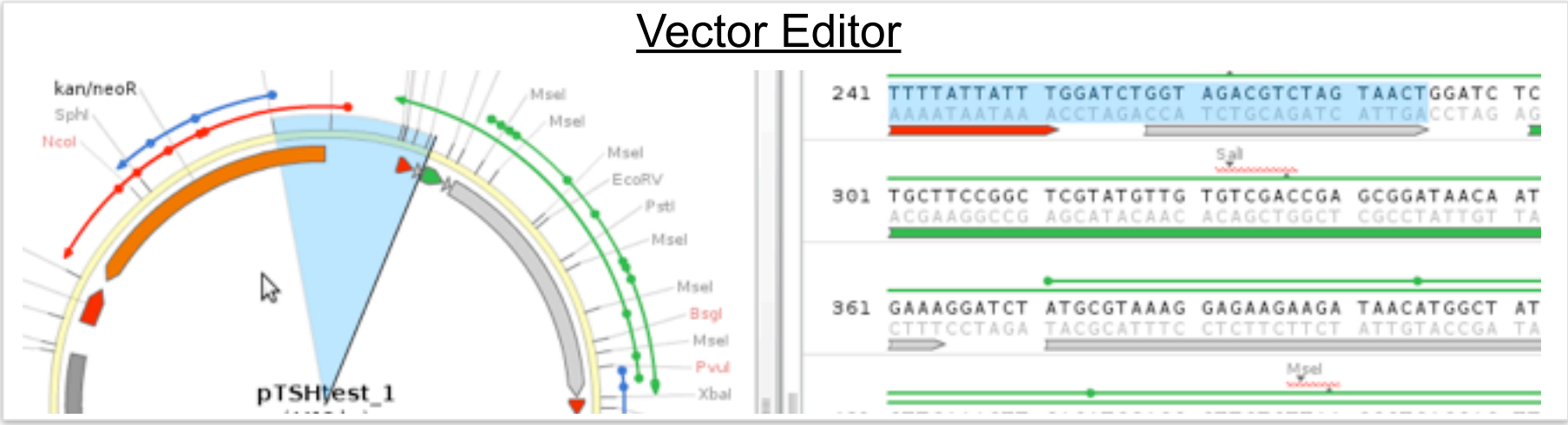
JBEI Registry of Biological Parts

- Enhanced feature annotation and sequence editing capabilities
- Open source and freely available
- JBEIR public registry website: https://public.jbeir.org

Sequence validation automation

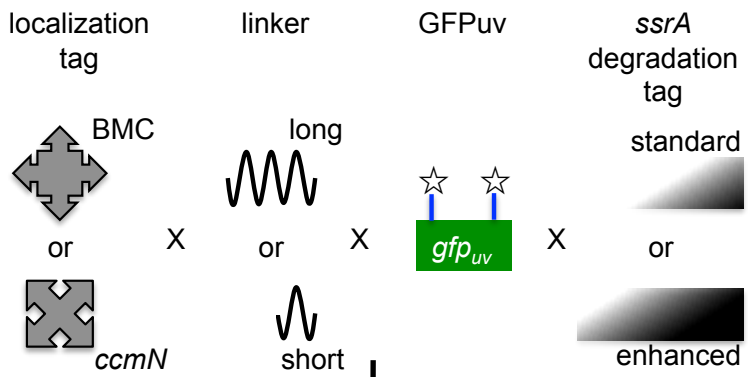


Vector Editor

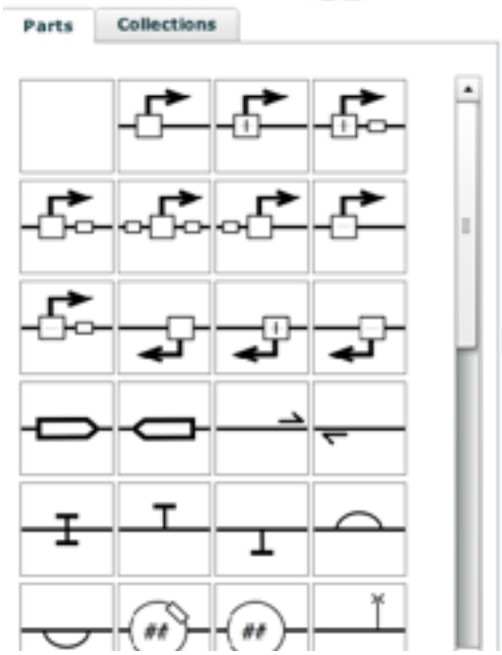


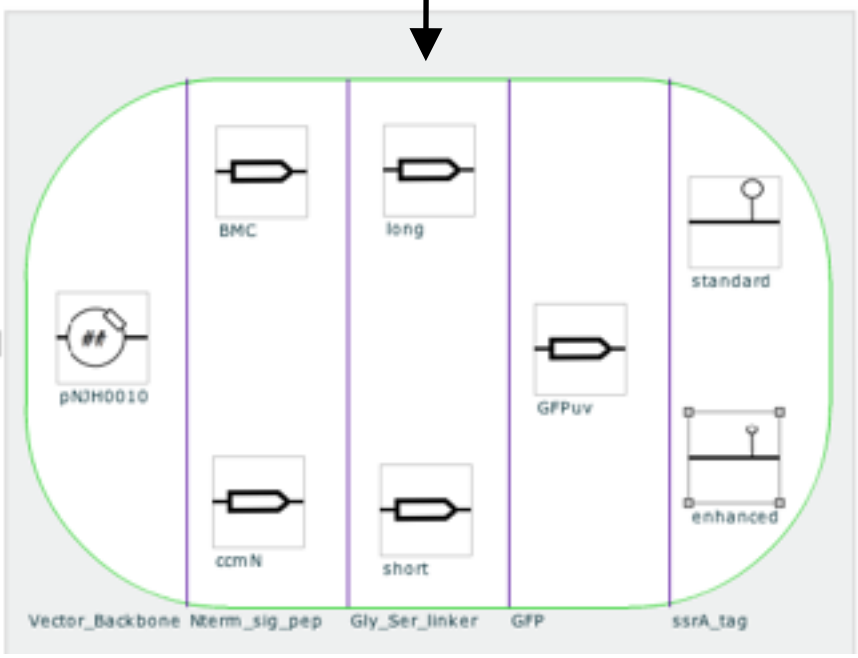
Device Editor: a BioCAD canvas

Visually represents and captures DNA assembly design schematic details



Douglas Densmore, Joanna Chen





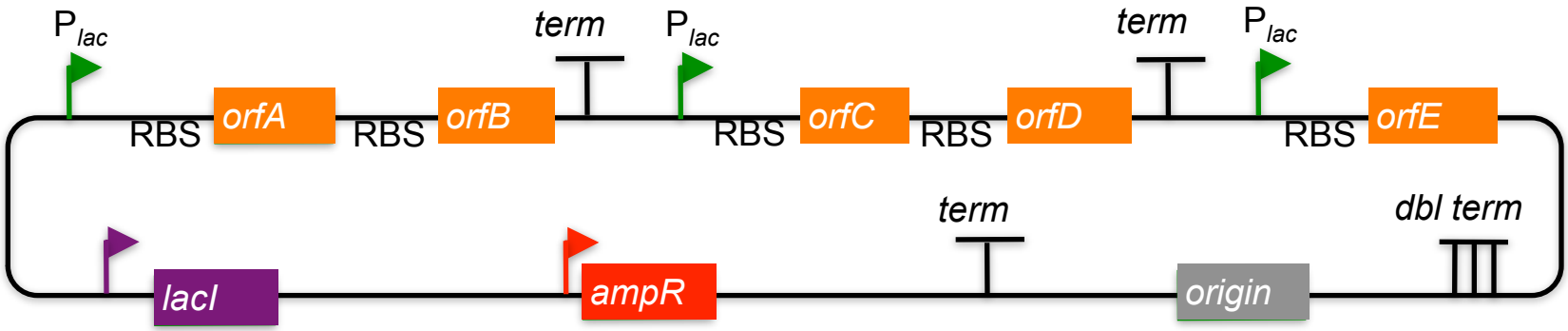
Device Editor: a BioCAD canvas

Facilitates the creation of Eugene design specification rules

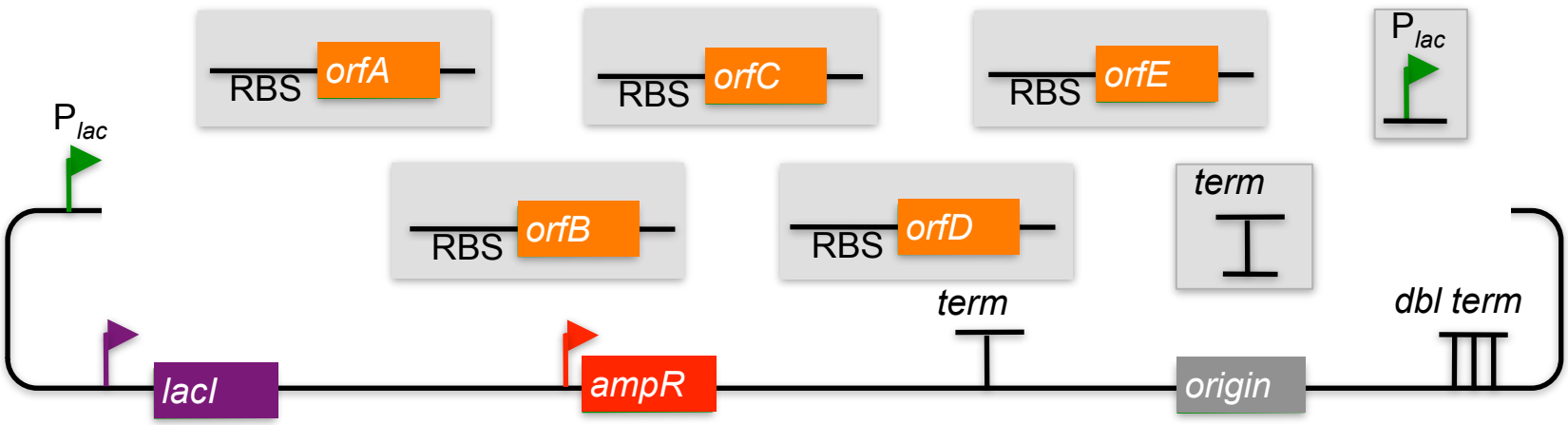
Eliminates undesired assembly combinations

The DNA assembly challenge

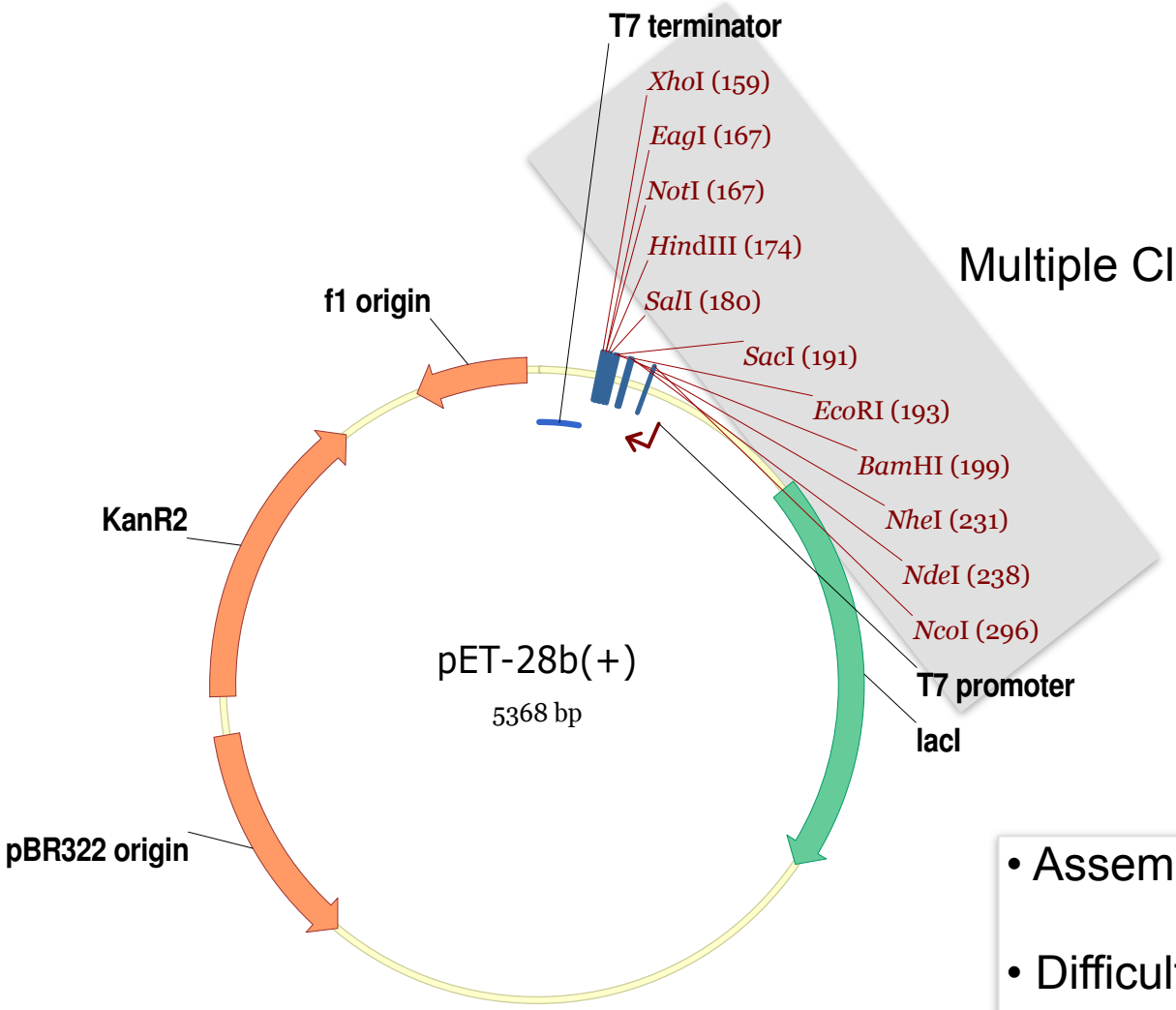
Metabolic pathway expression vector:



Composite biological parts:



The traditional approach

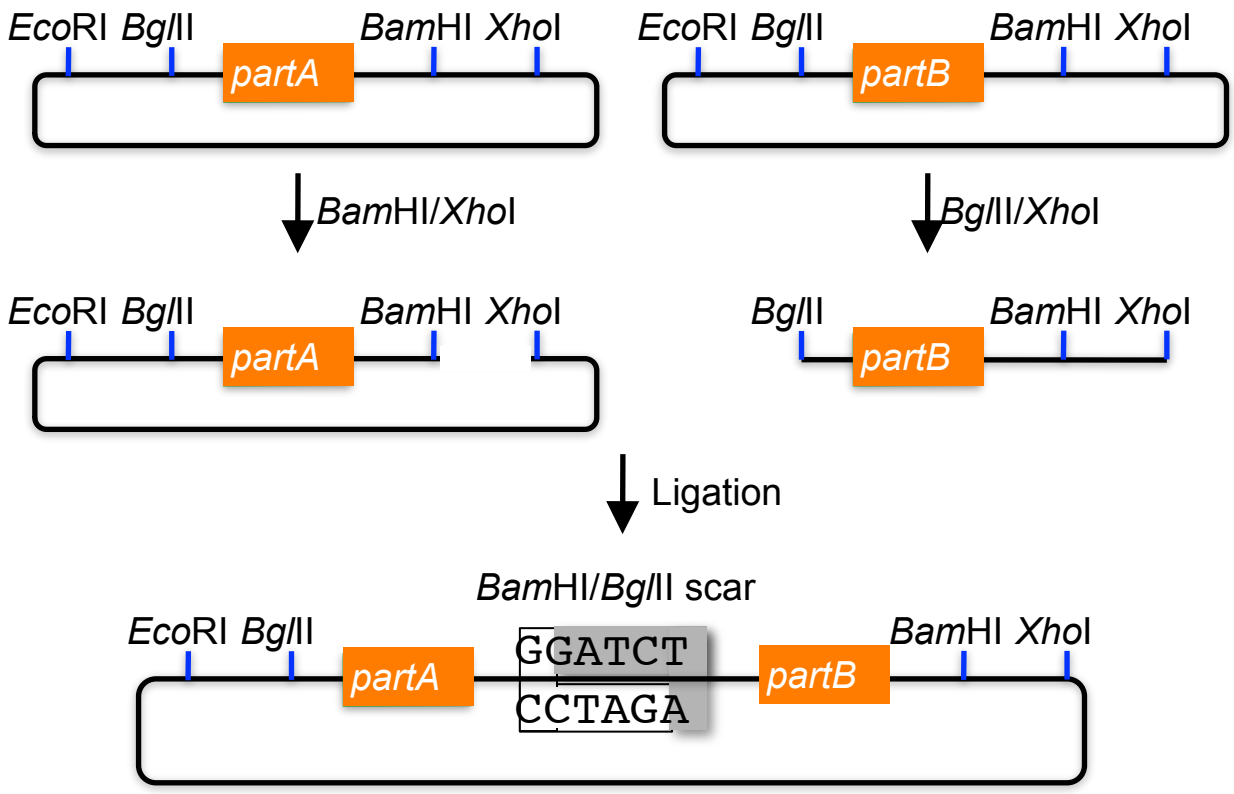


Multiple Cloning Site (MCS)

(EMD/Novagen)

- Assembly strategy not standardized
- Difficult to automate

The BioBrick approach



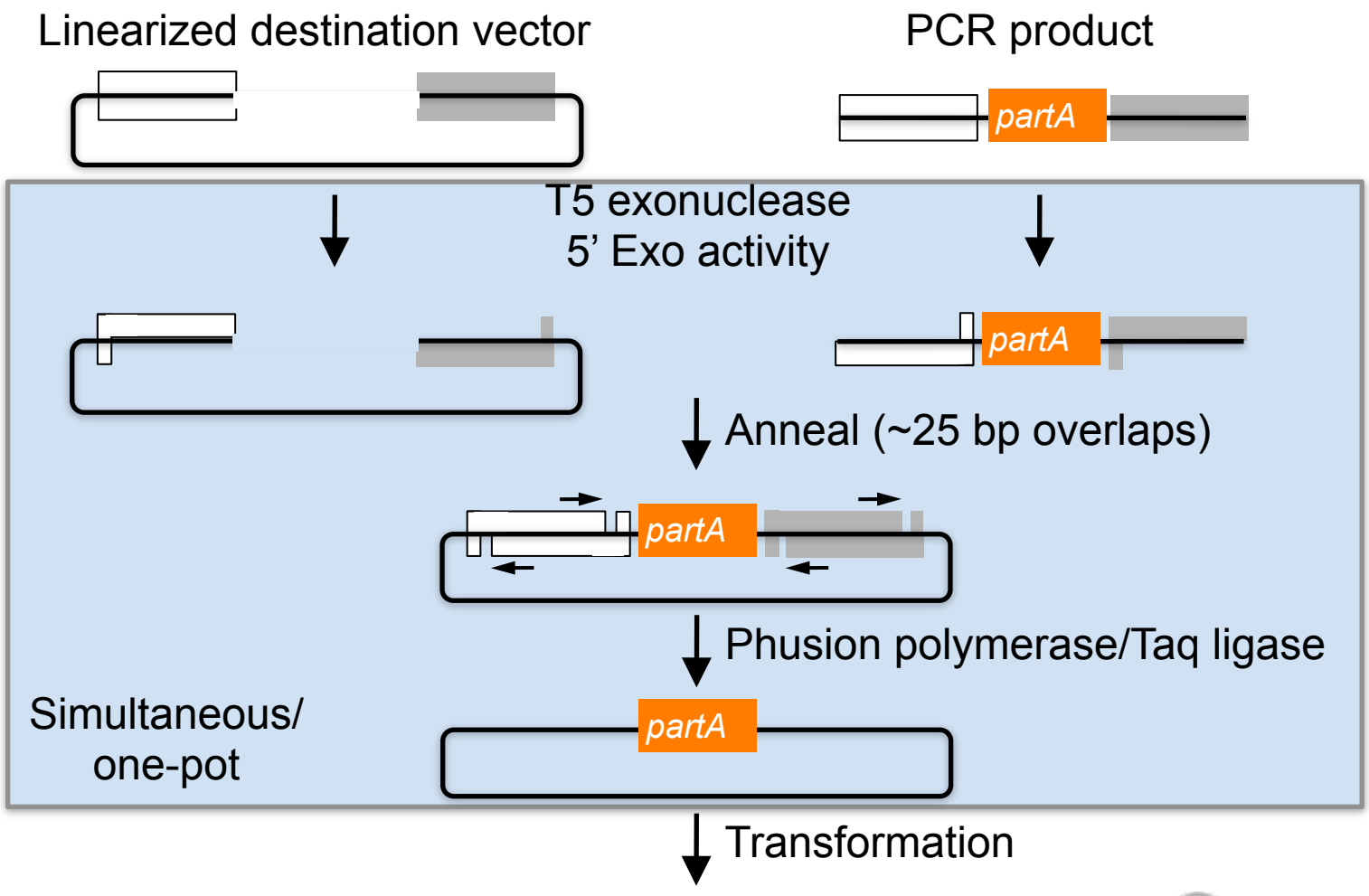
(Shetty *et al*. 2008, Anderson *et al*. 2010)

*Bam*HI	*Bgl*II
GGATCC	AGATCT
CCTAGG	TCTAGA

- Standardized assembly strategy
- Automatable (Anderson Lab, UC Berkeley)

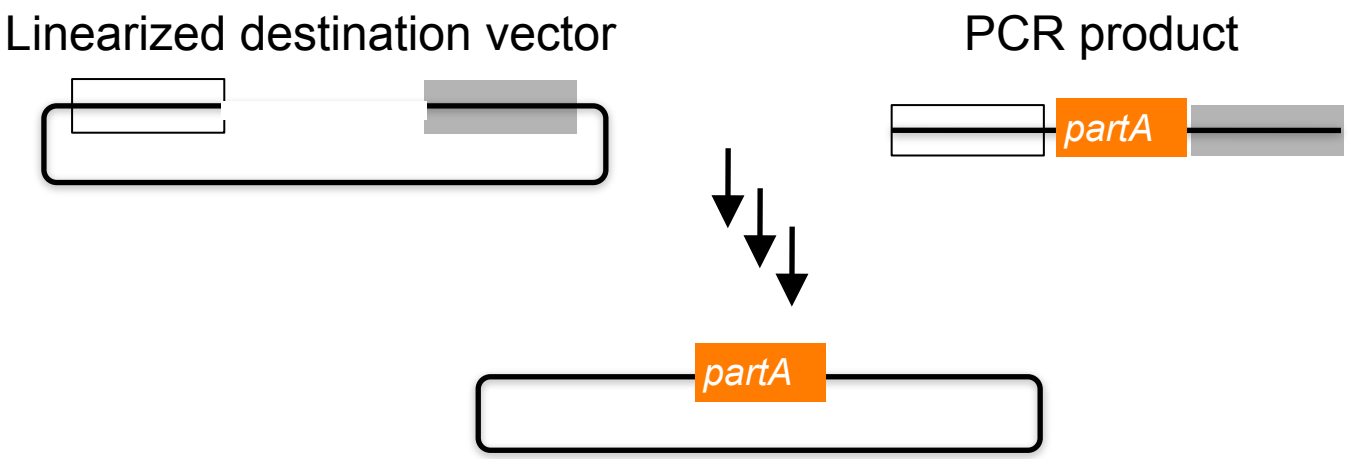
The Gibson approach

(Gibson *et al*. 2009)



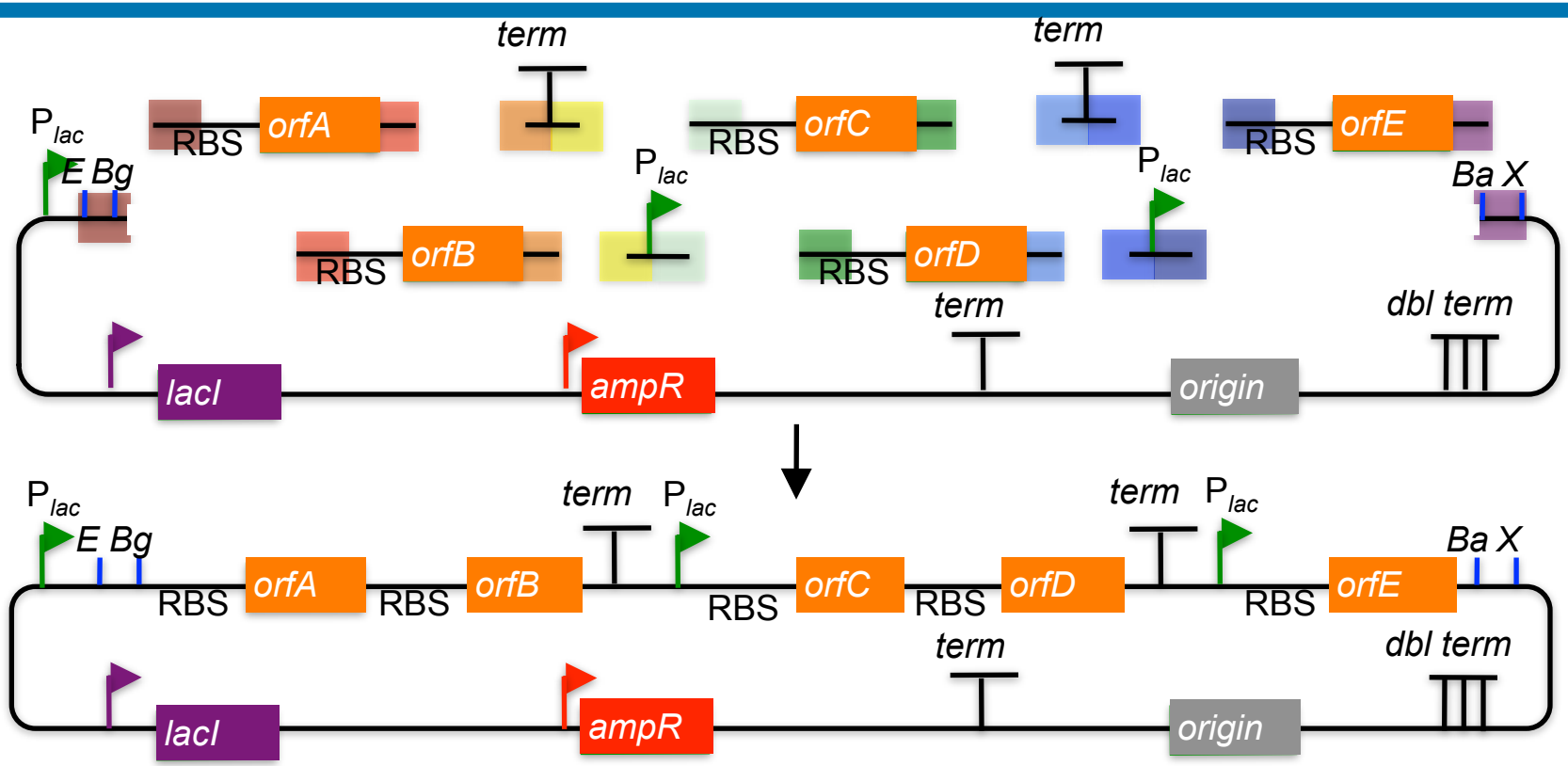
SLIC/Gibson/CPEC similarities

Identical starting materials and final assembled products



- Automatable assembly strategies
- Scar-less
- Destination vector independent (BioBrick compatible)

SLIC/Gibson/CPEC similarities



- One-pot concurrent multi-part assembly
- All parts accessible; combinatorial diversity preserved
- Must design oligos specifically for each assembly junction

j5: DNA assembly design

Given a list of biological parts to assemble, designs the most cost-effective strategy

Non-degenerate Part IDs and Sources							
Part			Location				
ID Number	Name	Source Plasmid	Reverse Complement	Start (bp)	End (bp)	Size (bp)	Sequence
0	pES119_SbfI_AscI_vector_fragment	pES119	FALSE	7965	4382	8529	CGCGCCT
1	atfA_3prime	pES119	FALSE	4383	4701	319	GGACTCA
2	BBa_B1002_term	BBa_B1002	FALSE	1	34	34	CGCAAAA
3	lacUV5_lacO	pBbE5a-RFP	FALSE	1449	1579	131	CCTGTGG
4	EcoRI_mutation	pES119	FALSE	4769	4769	1	A
5	pdc_without_3prime_end	pES119	FALSE	4769	6237	1469	AATTCAA
6	pdc_3prime_and_adhB	pES119	FALSE	6238	7704	1467	TGATGGT
7	BBa_B1006_term	BBa_B1006	FALSE	1	39	39	AAAAAAA
8	ITesA_5prime	pES119	FALSE	7770	7964	195	AATTCAA

Target Part Ordering/Selection/Strategy				
Part				
Order	ID Number	Name	Orientation	Strategy
0	0	pES119_SbfI_AscI_vector_fragment	forward	Linearized Destination Vector
1	1	atfA_3prime	forward	PCR
2	2	BBa_B1002_term	forward	Embed_in_primer_reverse
3	3	lacUV5_lacO	forward	SOE
4	4	EcoRI_mutation	forward	Embed_in_primer_reverse
5	5	pdc_without_3prime_end	forward	PCR
6	6	pdc_3prime_and_adhB	forward	PCR
7	7	BBa_B1006_term	forward	Embed_in_primer_reverse
8	3	lacUV5_lacO	forward	SOE
9	4	EcoRI_mutation	forward	Embed_in_primer_reverse
10	8	ITesA_5prime	forward	SOE

j5: DNA assembly design

Outputs oligo synthesis, PCR reaction, and assembly piece information

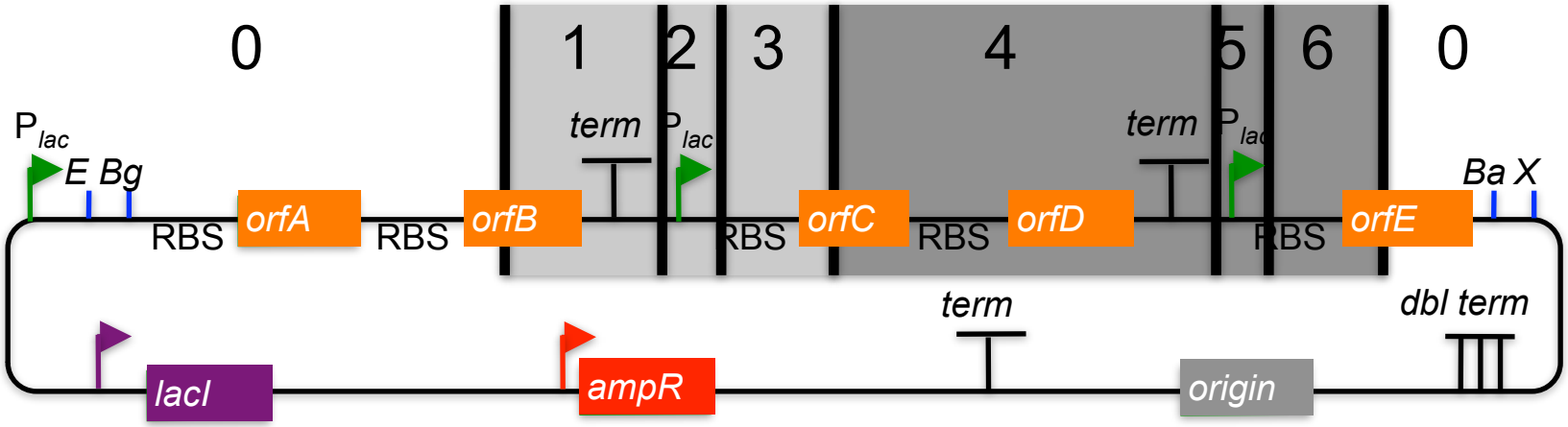
Oligo Synthesis					
ID Number	Name	Length	Tm	Cost ($)	Sequence
16	NJH00005_(lacUV5_lacO)_pure_forward	18	60.6	0	CCTGTGGC
17	NJH00017_(lacUV5_lacO)_forward	31	68.6	3.1	CGGGGTTT
18	NJH00007_(lacUV5_lacO)_pure_reverse	30	60.1	0	TGAAATTG
19	NJH00008_(lacUV5_lacO)_(EcoRI_mutation)_reverse	44	65.3	0	CTTCCTTT
20	NJH00018_(ITesA_5prime)_pure_forward	34	61.4	3.4	AATTCAAA
21	NJH00019_(ITesA_5prime)_forward	56	69.7	5.6	TGTGAGCC
22	NJH00020_(ITesA_5prime)_pure_reverse	23	61.9	2.3	CCAGTCCT
23	NJH00021_(ITesA_5prime)_reverse	53	79.2	5.3	GATGCTGT

PCR/SOE Reactions		Forward	Reverse Oligo					
ID Number	Template	Name	Name	Note	Mean Oligo Tm	ΔOligo Tm	Length	Amplified Sequence
0	pES119	NJH00001_	NJH00003_	PCR	53.4	4.7	319	GGACTCAACATAATT
1	pBbE5a-RFP	NJH00005_	NJH00007_	SOE	60.3	0.4	131	CCTGTGGCGCCGGTC
2	pES119	NJH00009_	NJH00011_	PCR	56.3	0.7	1469	AATTCAAAGGAAGGA
3	pES119	NJH00013_	NJH00015_	PCR	63.1	1.7	1467	TGATGGTCCGTACAA
4	pBbE5a-RFP	NJH00005_	NJH00007_	SOE	60.3	0.4	131	CCTGTGGCGCCGGTC
5	pES119	NJH00018_	NJH00020_	SOE	61.6	0.5	195	AATTCAAAGGAAGGA

Assembly Pieces (including the Linearized Destination Vector)								
ID Number	Type	Type ID	CPEC Tm Prev	CPEC Tm Next	Overlap Prev	Overlap Next	Length	Sequence
0	Linearized Destination Vector	0	67.7	63.4	30	26	8529	CGCGCCTT
1	PCR	6	61.4	70.0	26	26	392	ATAGTGCT
2	SOE	7	70.0	49.9	26	35	158	GGGGTTTT
3	PCR	8	60.5	60.6	35	32	1510	TGTGAGCC
4	PCR	9	56.5	66.7	32	26	1532	AAGTTATG
5	SOE	10	66.7	49.9	26	35	158	CGGGGTTT
6	SOE	11	60.5	69.9	35	30	247	TGTGAGCC

j5: DNA assembly design

Identifies assembly piece incompatibilities and suggests hierarchical assembly



Incompatibilities between Assembly Pieces:

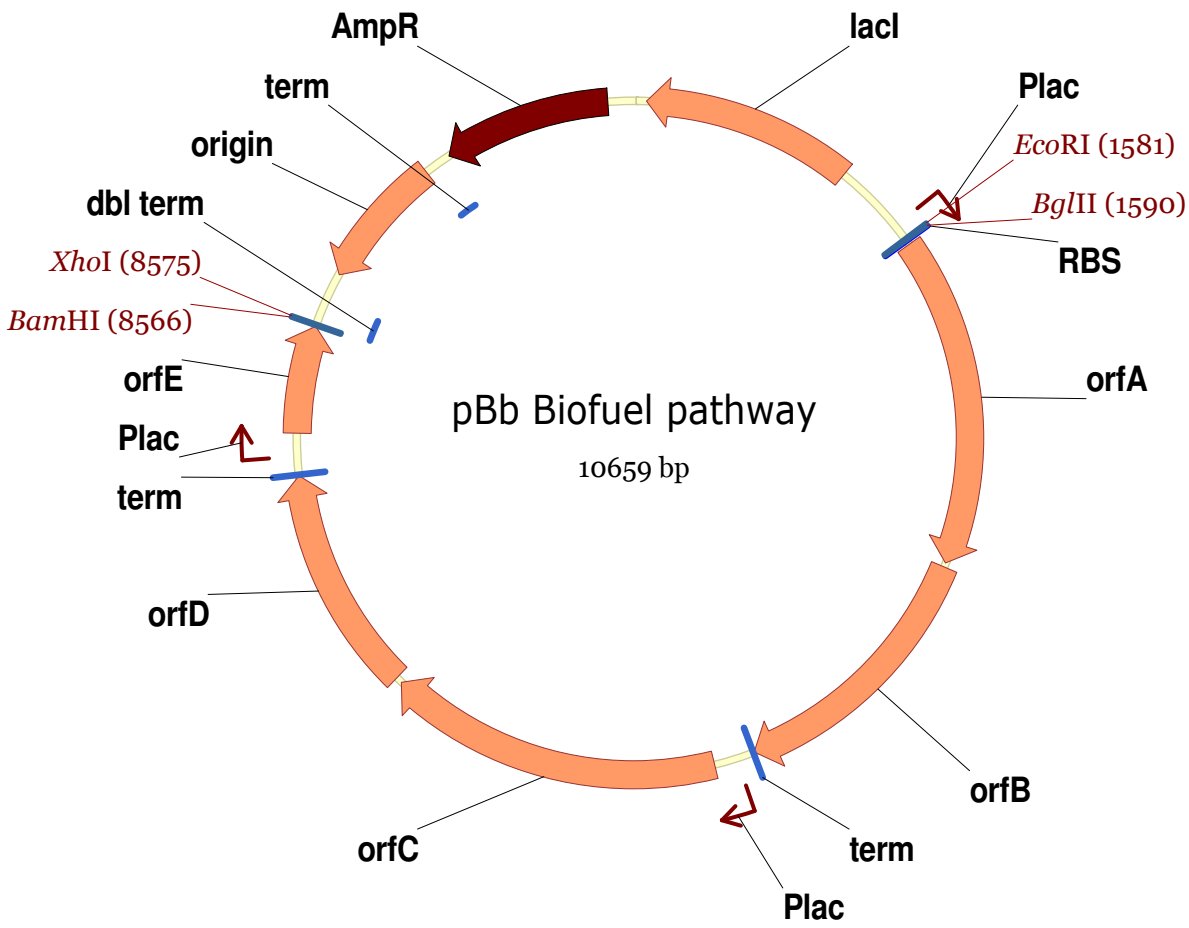
Assembly Piece	Left End	Right End
0	NONE	NONE
1	NONE	0, 4, 5
2	NONE	5, 6
3	0, 5, 6	NONE
4	NONE	0, 2
5	NONE	2, 3
6	0, 2, 3	NONE

Suggested Assembly Piece Contigs For Hierarchical Assembly:

Contig	Assembly Pieces Contained
0	0
1	1, 2, 3
2	4, 5, 6

j5: DNA assembly design

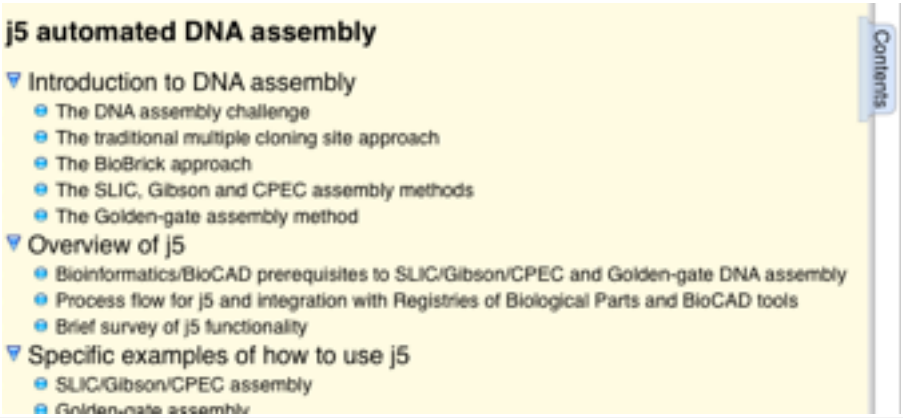
Generates an annotated genbank sequence file for the resulting assembly



j5 and DeviceEditor online

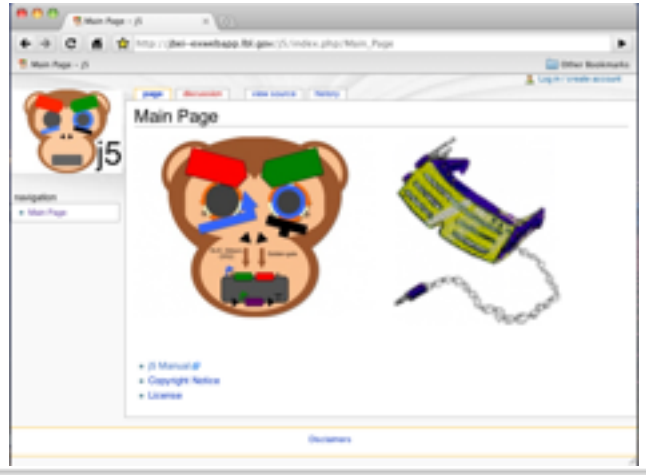
j5 web interface

j5 user's manual

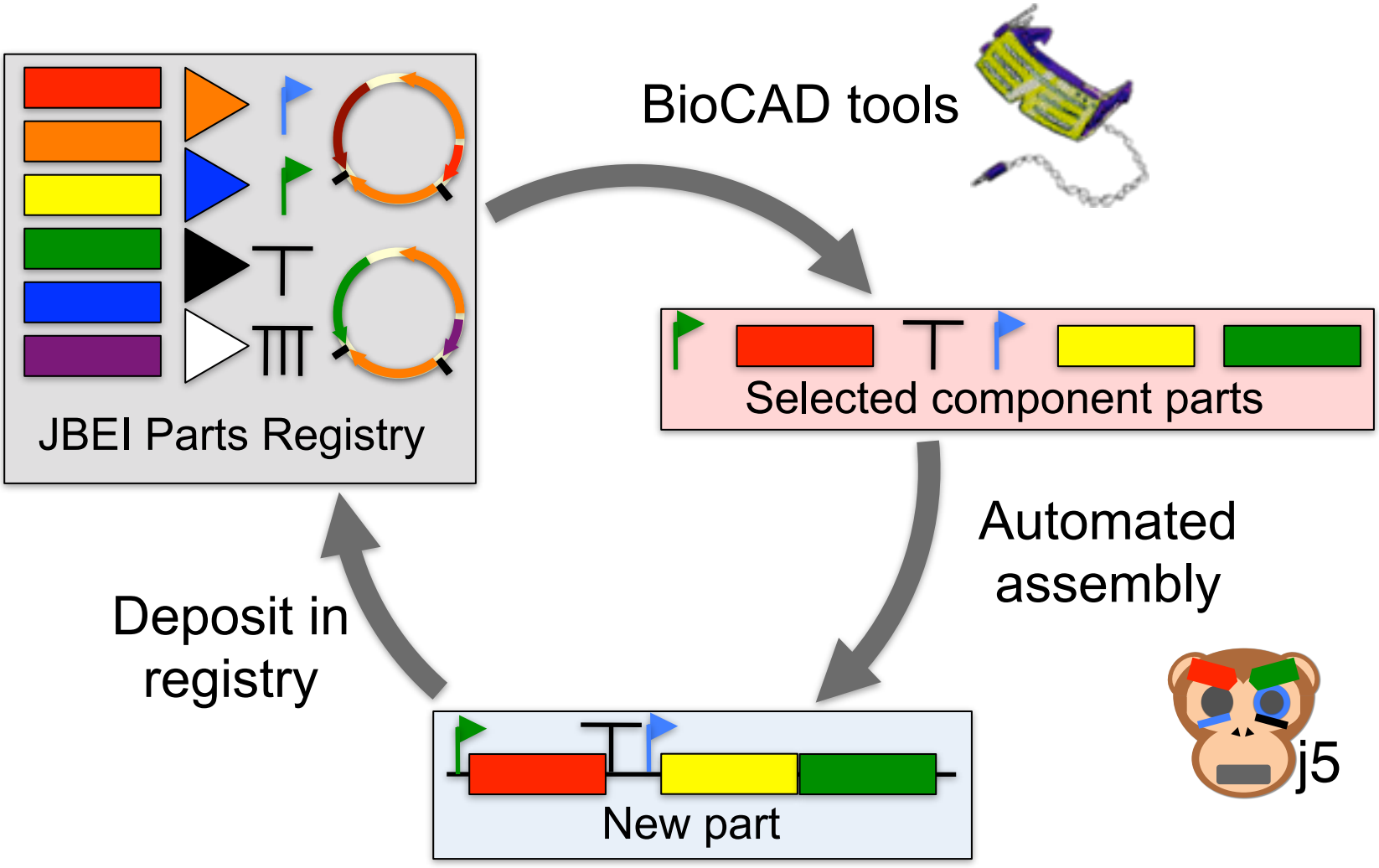


- Available at no cost to academic, non-profit, and government labs
- Commercial licensing and 14-day trial periods available
- Contact TTD@lbl.gov for more information

http://jbei-exwebapp.lbl.gov/j5



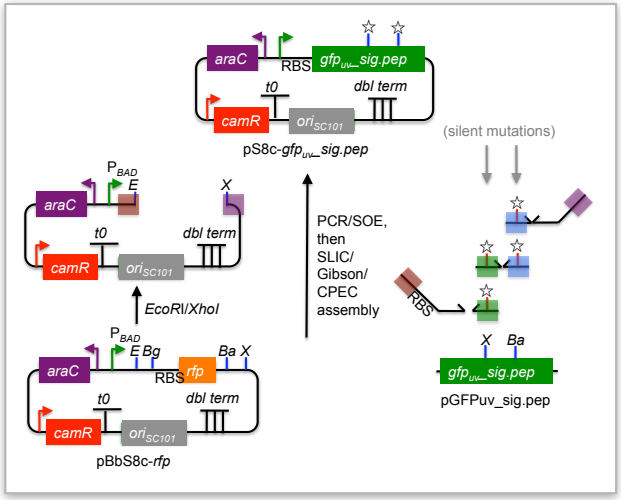
Synthetic Biology – an integrated process



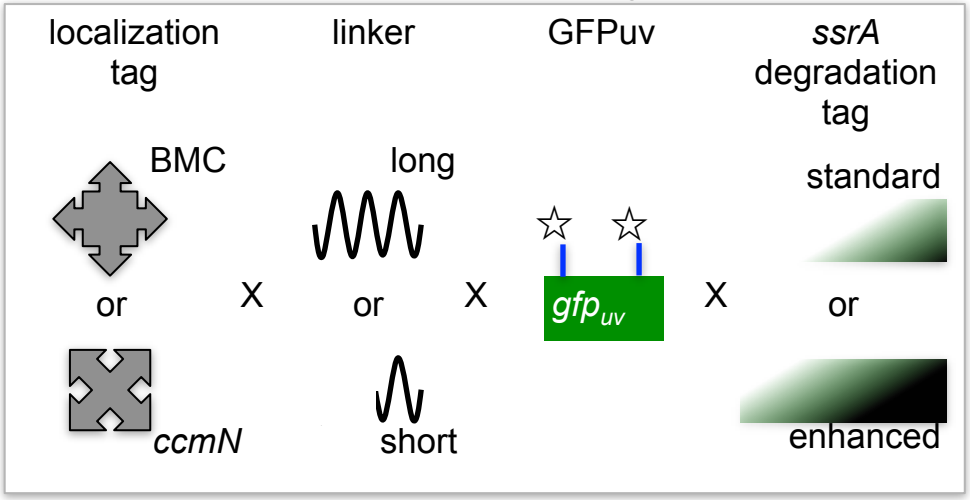
Experimental efforts underway

Execution and validation of j5 DNA assembly designs

SLIC/Gibson/CPEC

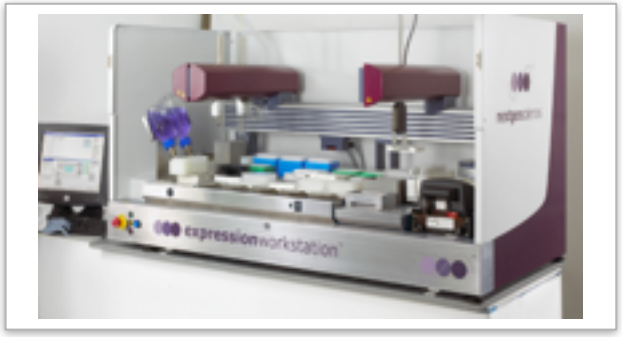


Combinatorial Golden-gate

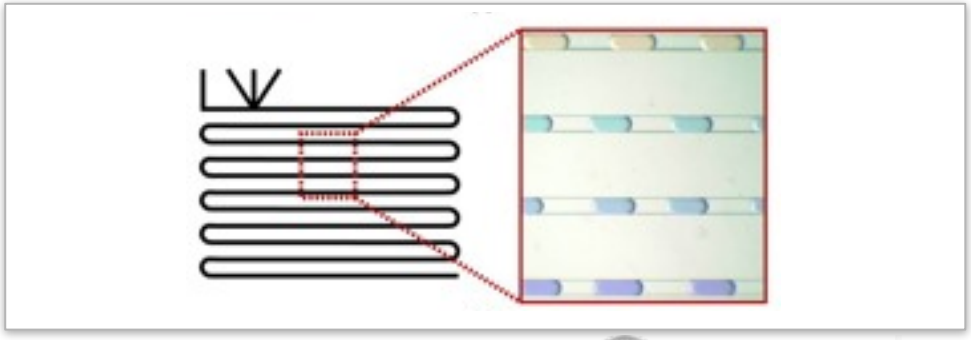


Integration of high-throughput devices into DNA assembly workflows

Liquid-handling robotics



Droplet-based microfluidic chips (Gerdts *et al.*, 2008)



Acknowledgements

JBEI Registry
- Timothy Ham
- Zinovii Dmytriv

DeviceEditor
- Douglas Densmore
- Joanna Chen

j5-designed assembly validation
- John W. Thorne
- Jay Kinney, Anna Chen
- Rafael Rosengarten

Robotics discussions
- Masood Hadi

Microfluidics discussions
- Anup Singh
- Rajiv Bharadwaj
- Aarthi Chandrasekaran
- Chieh Chang

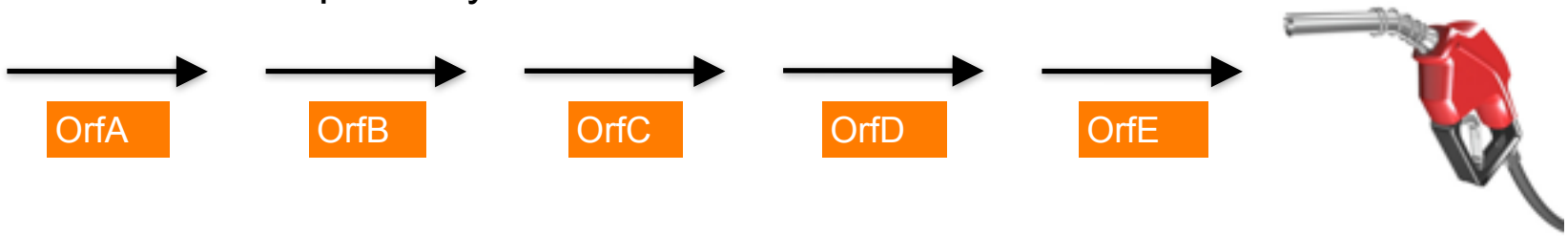
Feedback and discussions
- Jay Keasling
- Keasling Lab
- JBEI Scientific Divisions

Operations support
- JBEI staff

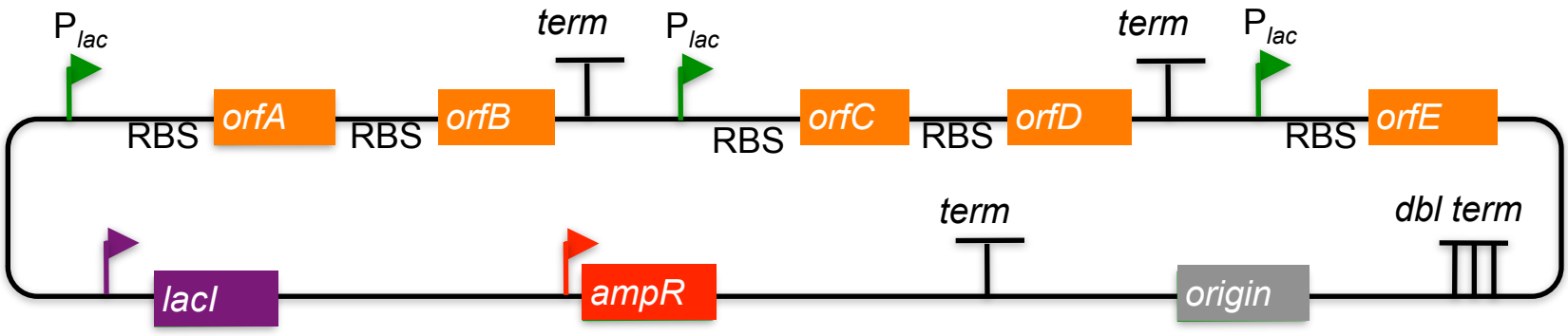
Funding
- DOE Genomics/GTL/BRC (JBEI)
- Berkeley Lab LDRD

The DNA assembly goal

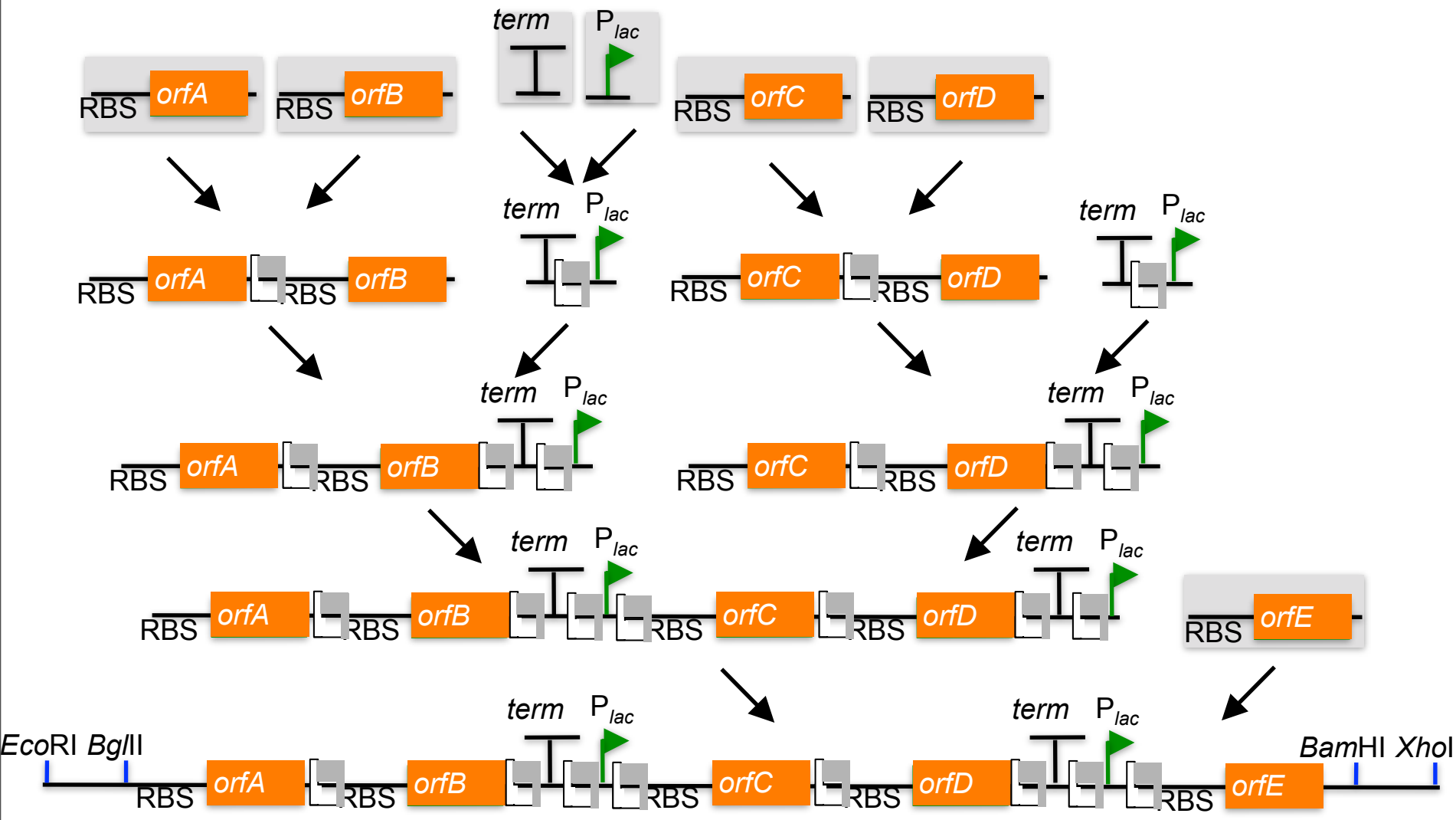
Biofuel metabolic pathway:



Metabolic pathway expression vector:

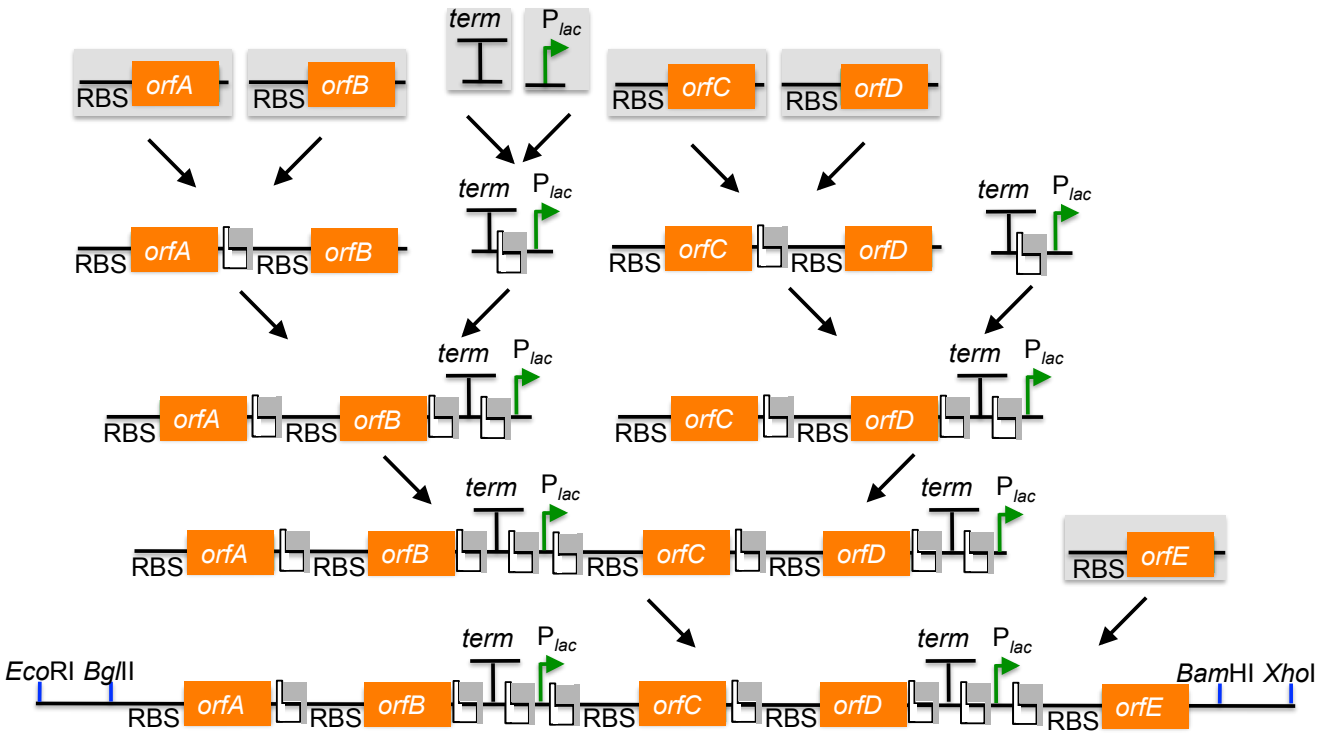


BioBrick parallel assembly



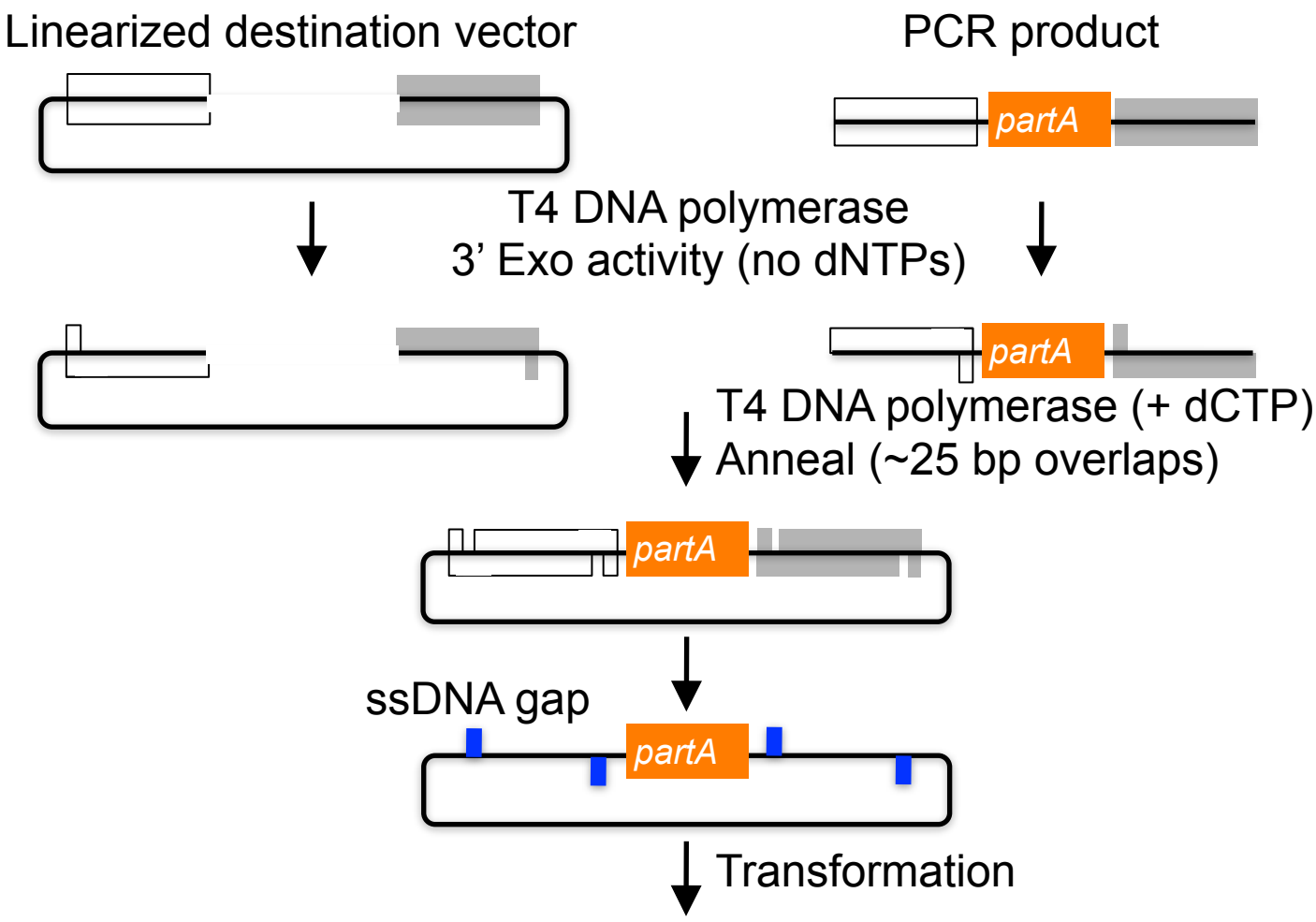
BioBrick assembly considerations

- Introduction of scar sites
- After assembly: difficult to modify central parts
- Potential combinatorial diversity reduction



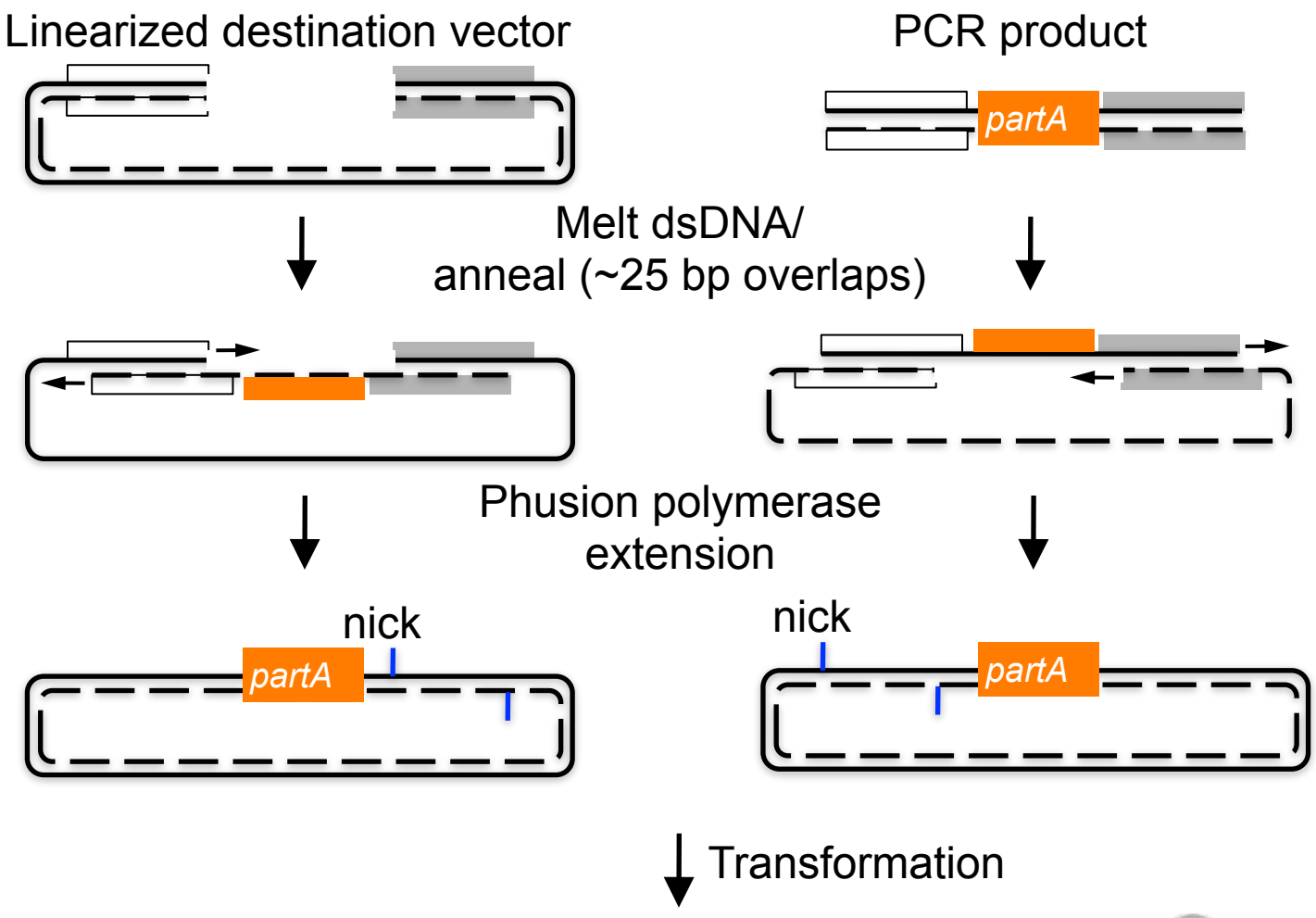
The SLIC approach

Sequence and Ligation-Independent Cloning (Li *et al*. 2007)



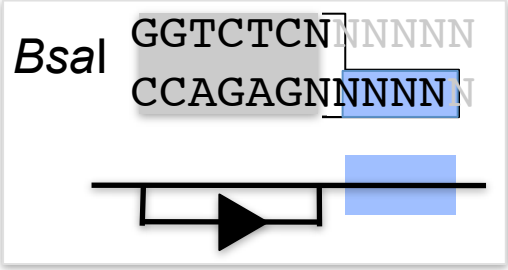
The CPEC approach

Circular Polymerase Extension Cloning (Quan *et al*. 2009)



The Golden Gate approach

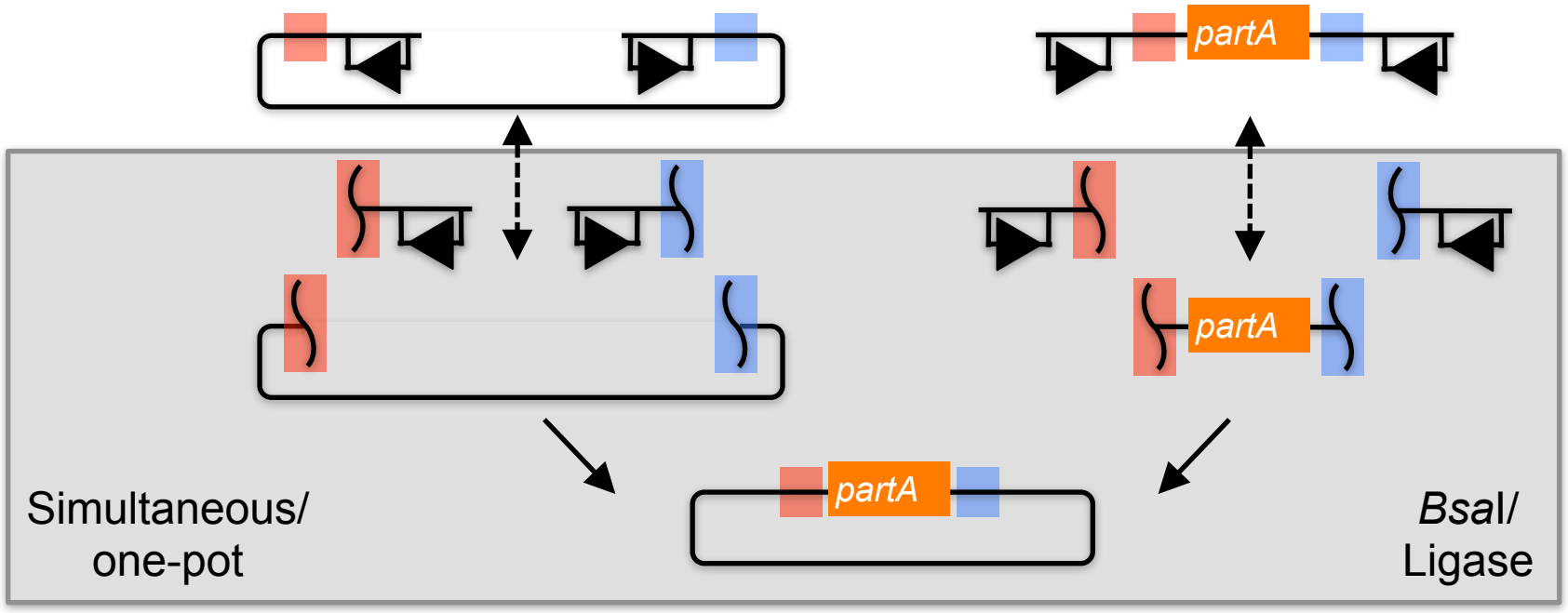
A variation of
(Engler *et al*. 2008,
Engler *et al*. 2009)



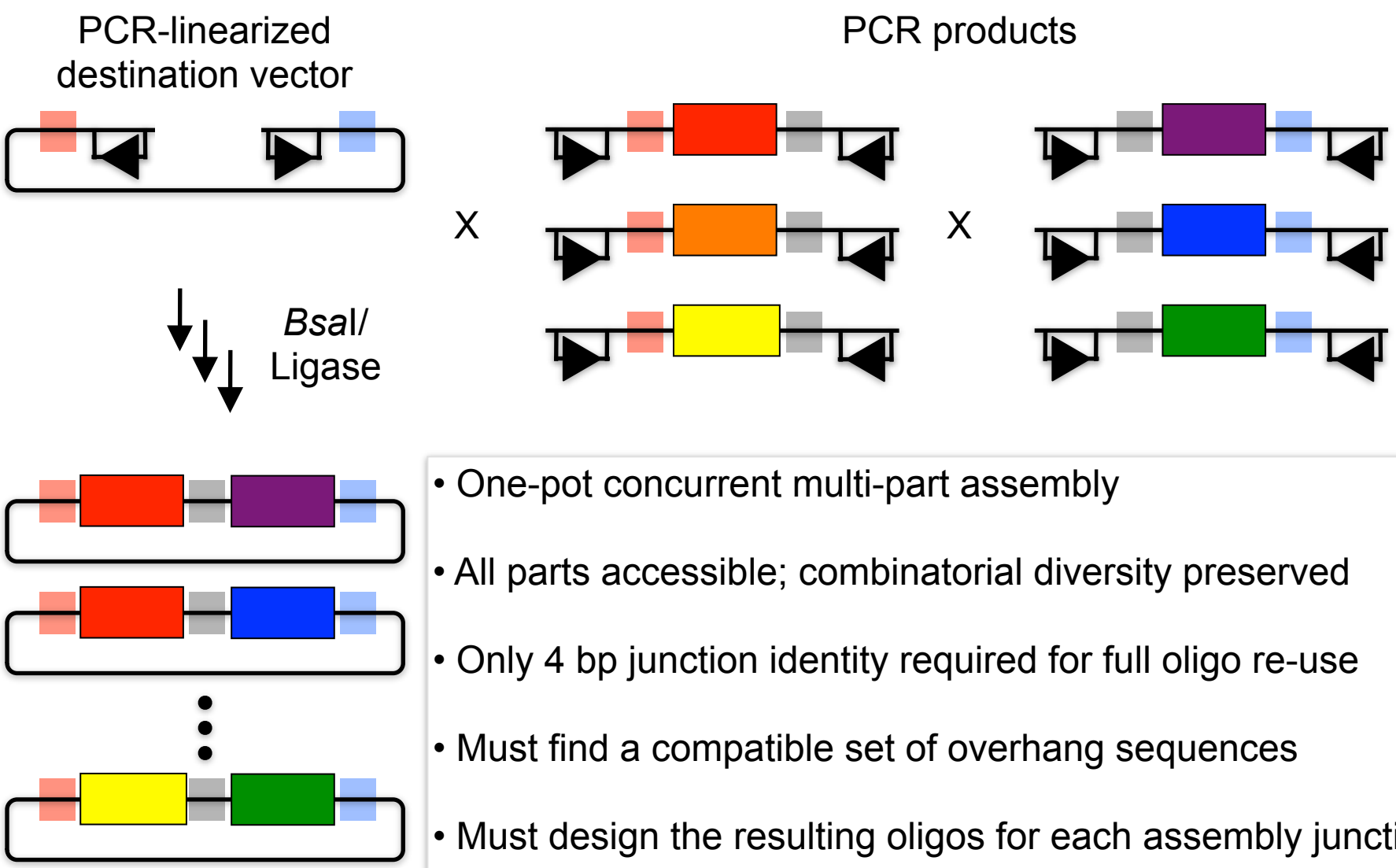
- Automatable assembly strategy
- Scar-less
- Largely vector independent

PCR-linearized destination vector

PCR product



Combinatorial Golden Gate assembly



- One-pot concurrent multi-part assembly
- All parts accessible; combinatorial diversity preserved
- Only 4 bp junction identity required for full oligo re-use
- Must find a compatible set of overhang sequences
- Must design the resulting oligos for each assembly junction

j5: DNA assembly design

Outputs oligo synthesis, PCR reaction, and assembly piece information

Oligo Synthesis					
ID Number	Name	Length	Tm	Cost ($)	Sequence
4	NJH00146_(atfA_3prime)_pure_forward	26	55.7	0	GGACTCAACATAATTTCT
5	NJH00166_(atfA_3prime)_forward	41	65.5	4.1	TTTTTTGGTCTCACAGGA
6	NJH00148_(atfA_3prime)_pure_reverse	36	59.7	0	ATCCTTAATTGGCTGTTT
7	NJH00167_(atfA_3prime)_(BBa_B1002_term)_reverse	85	76.8	48.5	TTTTTTGGTCTCAGGGCG
8	NJH00025_(lacUV5_lacO)_pure_forward	18	60.6	0	CCTGTGGCGCCGGTGAT
9	NJH00168_(lacUV5_lacO)_forward	33	70.0	3.3	TTTTTTGGTCTCAGCCCT

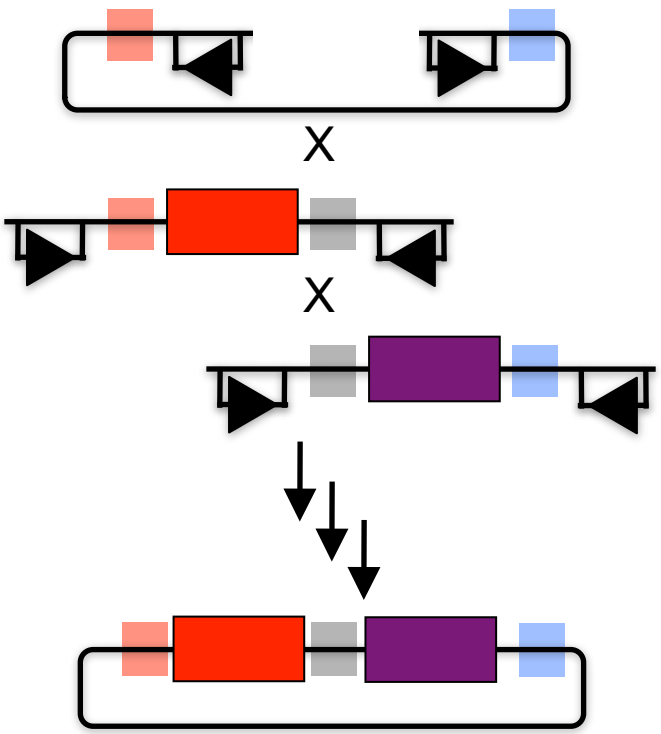
PCR Reactions		Forward	Reverse				
ID Number	Template	Oligo	Oligo	Mean Oligo Tm	Δ Oligo Tm	Length	Amplified Sequence
0	pES119	NJH00162_	NJH00164_	62.3	2.5	8529	CGCGCCTTCCGGCTCTG
1	pES119	NJH00146_	NJH00148_	57.7	4.0	319	GGACTCAACATAATTTCT
2	pBbE5a-RFP	NJH00025_	NJH00094_	60.3	0.4	131	CCTGTGGCGCCGGTGAT
7	PCR Reaction 0	NJH00163_	NJH00165_	71.8	4.4	8563	TTTTTTGGTCTCAGACTG
8	PCR Reaction 1	NJH00166_	NJH00167_	71.1	11.4	383	TTTTTTGGTCTCACAGGA
9	PCR Reaction 2	NJH00168_	NJH00169_	68.1	3.9	162	TTTTTTGGTCTCAGCCCT

Assembly Pieces (including the Linearized Destination Vector)							
ID Number	Type	Type ID	5' Overhang	3' Overhang	Relative Overhang	Length	Sequence
0	PCR Linearized Destination Vector	7	GACT	CAGG	0	8563	TTTTTTGGTCTCAGACTG
1	PCR	8	CAGG	GCCC	0	383	TTTTTTGGTCTCACAGGA
2	PCR	9	GCCC	AAAA	0	162	TTTTTTGGTCTCAGCCCT
3	PCR	10	AAAA	CCAT	-1	1498	TTTTTTGGTCTCAAAAAT
4	PCR	11	CCAT	TTCC	0	1537	TTTTTTGGTCTCACCATG
5	PCR	12	TTCC	TCAA	-2	160	TTTTTTGGTCTCATTCCT
6	PCR	13	TCAA	GACT	-4	223	TTTTTTGGTCTCATCAAA

j5: DNA assembly design

Exhaustively searches for the optimal set of compatible overhang sequences

Compatible sets of overhang sequences



5′-TGCCTACC…
ATGG…
…AGTT
…TCAAACGG-5′

Desired overhang annealing

5′-TGCCTACC…
ATGG…
…TAGC
…ATCGGCAA-5′

No overhang annealing

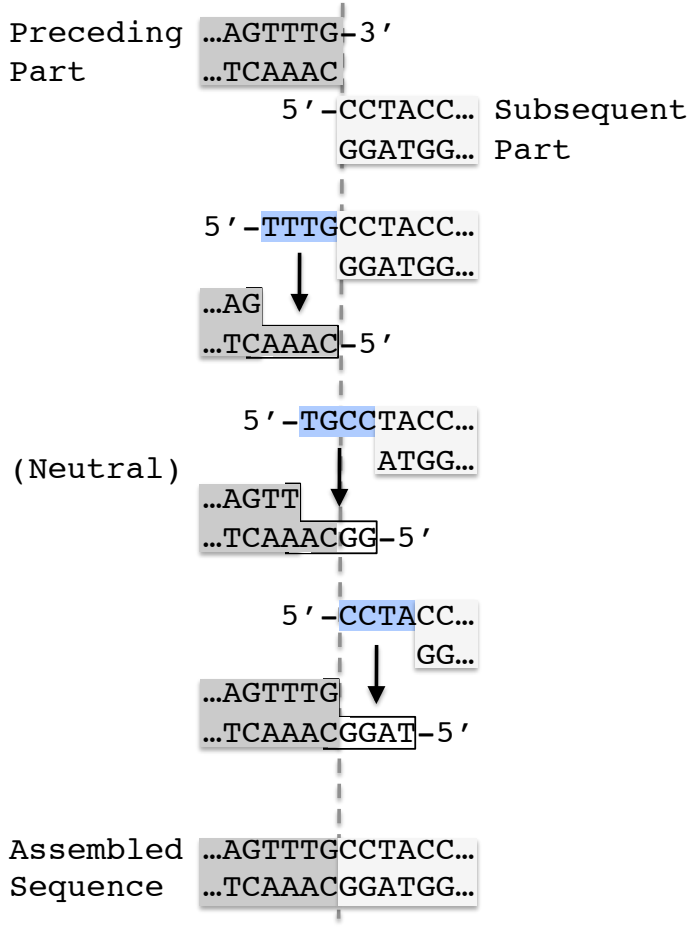
5′-TGCCTACC…
ATGG…
…TATT
…ATAAAGGG-5′

Undesired overhang annealing

j5: DNA assembly design

Exhaustively searches for the optimal set of compatible overhang sequences

Golden Gate overhang sequence selection options



Device Editor: a BioCAD canvas

Generates j5 input files and controls j5 parameters

J5 Files | J5 Properties

J5 Sequence File Generate | Genbank File Create | Download

J5 Parts File Generate | FASTA File Create

J5 Target Parts File Generate

J5 Eugene Rules File Generate

J5 Files | J5 Properties

Master Oligo Num of Digits (5)	5	Primer GC Clamp (2)	2
Master Plasmid Num of Digits (5)	5	Primer Min Size (18)	18
Gibson Overlap BPs (26)	26	Primer Max Size (36)	36
Maximum Num of Oligo BPs (110)	110	Primer Min TM (60)	60
Minimum Gibson Fragment Size BPs (250)	250	Primer Max TM (70)	70
Golden Gate Overhang BPs (4)	4	Primer Max TM Diff (5)	5
Golden Gate Recognition Sequence (GGTCTC)	GGTCTC	Primer TM Santa Lucia (1)	1
Golden Gate Term Extra Seq	TTTTTTGGTCTCA	Primer Salt Corrections (1)	1
Max Identities Golden Gate Overhangs (2)	2	Primer DNA Conc (250)	250
Oligo Synthesis Cost Per BP (US) (0.1)	0.1	MisPriming 3Prime Boundary (4)	4
Oligo Purification Cost Per Piece (US) (40)	40	MisPriming Min TM (45)	45
Oligo Max Length No Page Purification (60)	60	MisPriming Salt Conc (0.05)	0.05
Minimum PCR Product BPs (100)	100	MisPriming Oligo Conc (2.5e-7)	2.5e-7
Direct Synthesis Cost Per BP (US) (0.39)	0.39	Output Sequence Format (Genbank)	Genbank
Direct Synthesis Min. Cost Per Piece (US) (159)	159		

Write to File | Return to Defaults | Quit Without Saving | *(Default Value)