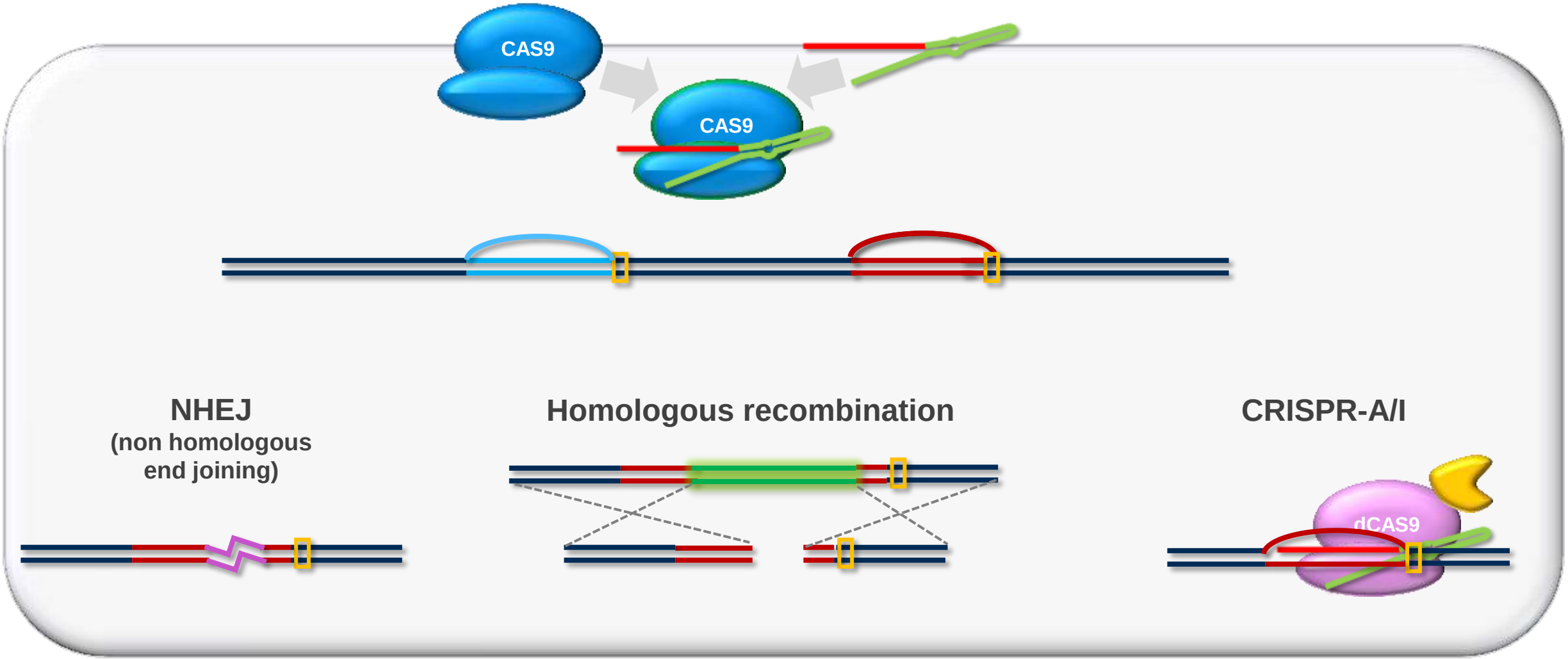
A CRISPer Guide for Generating Superior sgRNA Libraries

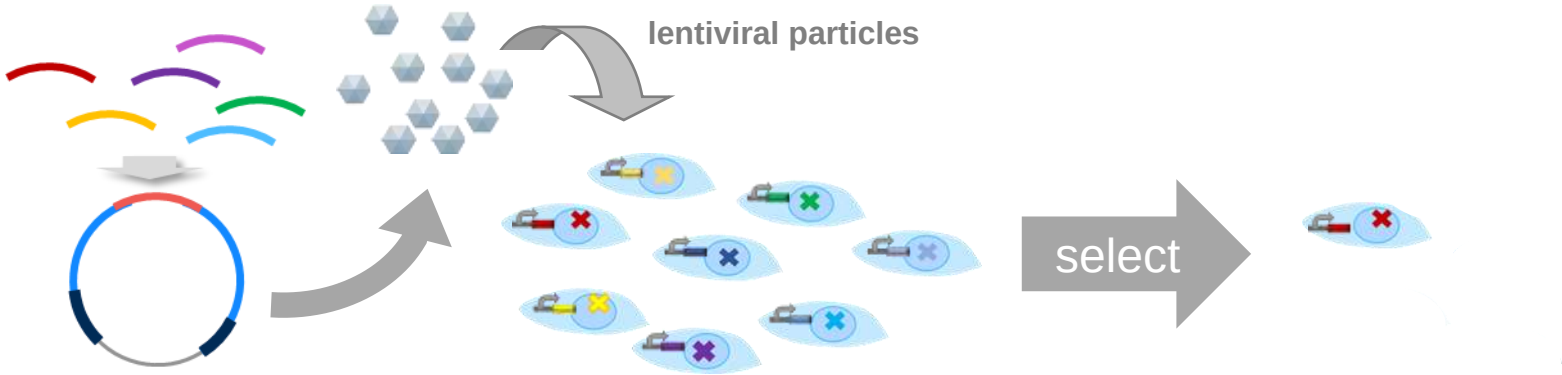
Carsten Carstens

Senior Scientist, R&D

Agilent Technologies


CAS9
CAS9
NHEJ
(non homologous
end joining)
Homologous recombination
CRISPR-A/I
dCAS9

Functional screening using CRISPR/CAS

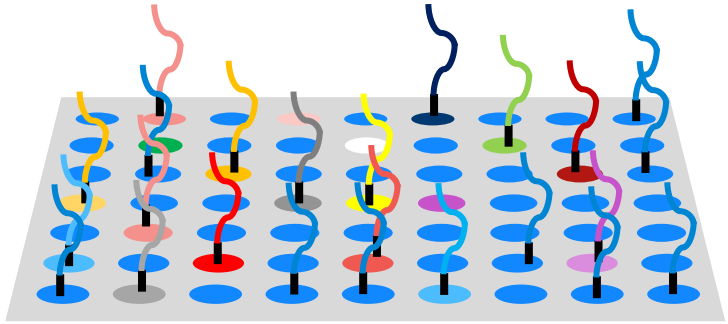


Highly parallel
DNA synthesis
capabilities

Next generation
sequencing

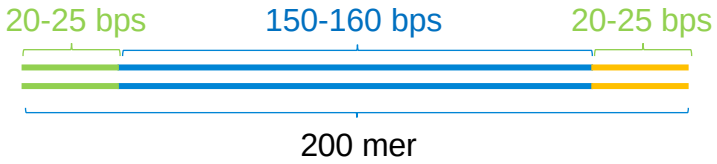
Statistical analysis packages

- RIGER
- HiT Select
  Diaz et al., *NAR* (2014)
- MAGeCK
  Li et al., *Genome Biology* **15**:554 (2014)
- casTLE
  Morges et al., *Nature Biotech* ePub April 11th (2016)



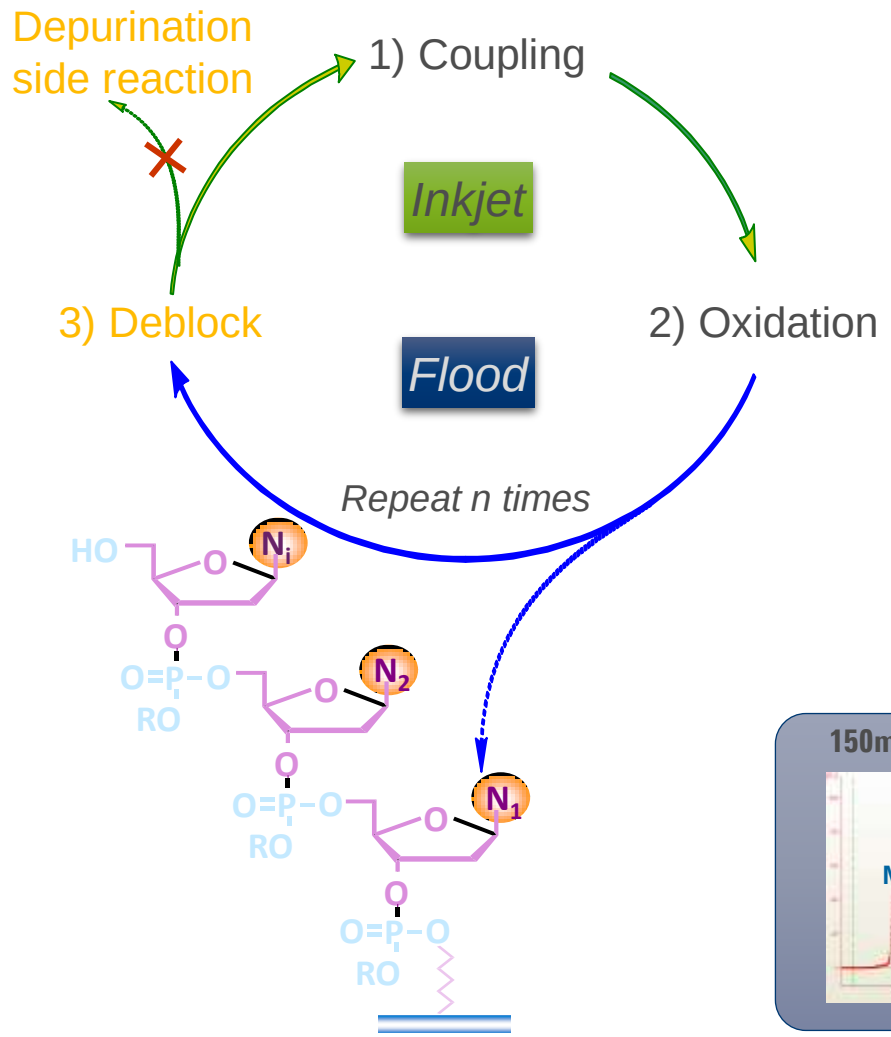
OLS libraries:
Up to 120,000 user-definable sequences /chip

Due to the limited amount synthesized on single features, libraries need to be amplified prior to use

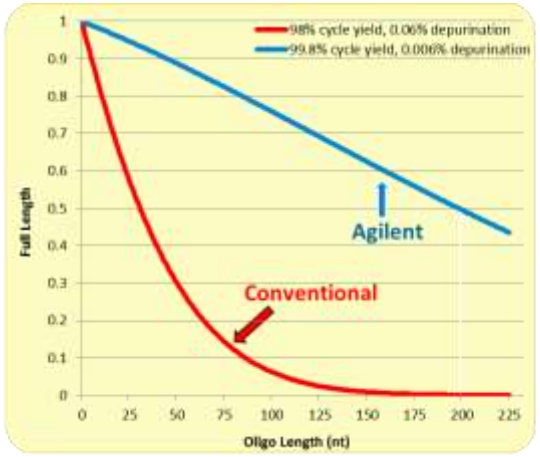


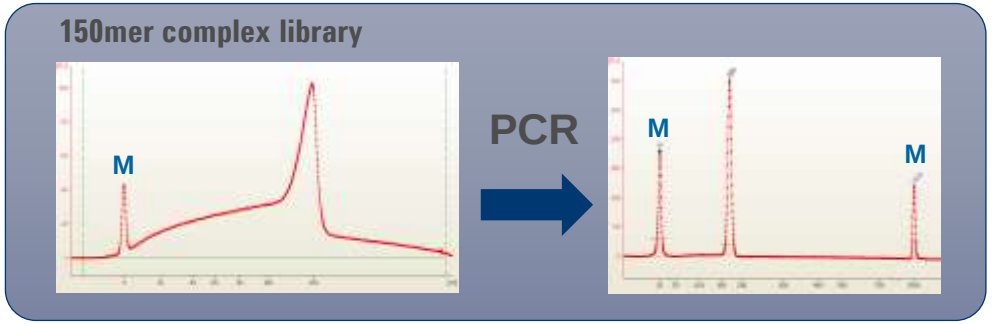
- Libraries are provided as dsDNA
- Several sub-libraries can be printed on the same chip
- PCR acts as clean-up step
- Control features can be integrated into the chip design

Chemical Synthesis: Achieving High synthesis efficiency



Long length synthesis is achieved
by improved cycle yield
- ↑ coupling efficiency
- ↓ depurination
- ↑ consistency





Applications for OLS libraries

Gene and genome synthesis

Selective target enrichment for high throughput sequencing (HTS)

Probes for *in situ* hybridizations

High throughput site directed mutagenesis

Plasmid based libraries of short sequences

GEN9

- SureSelect
- HaloPLex

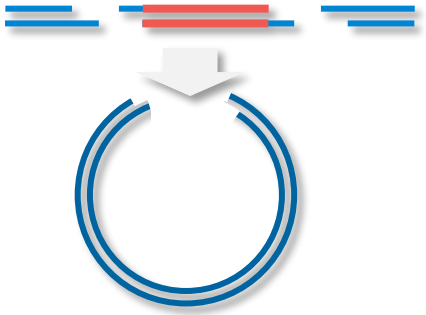
SureFISH

QuikChange-HT

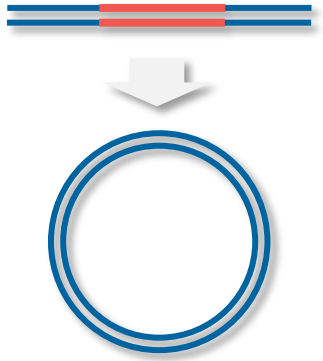
For Research Use Only. Not for use in diagnostic procedures

Cloning an oligo library into a plasmid vector

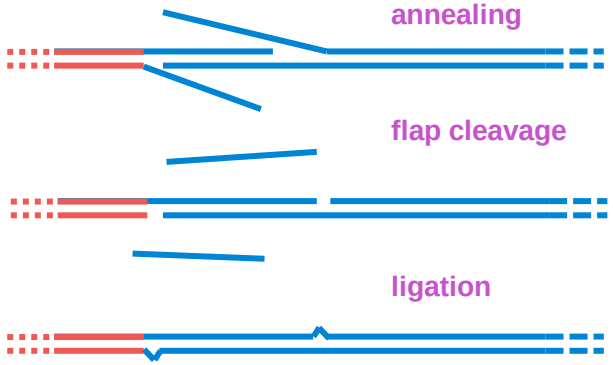
restriction/ligation



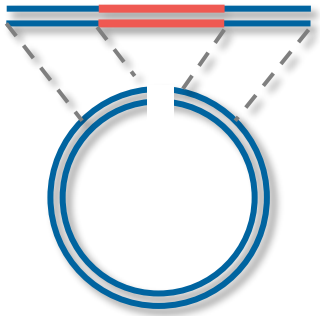
- Requires processing of OLS (unless Type IIS enzymes are used).
- Library members can't contain restriction site



Novel overlap-based assembly method using flap cleavage-mediated strand joining



Overlap-based assembly



- No processing required
- No restrictions on content
- Seamless integration

Constructing plasmid libraries from amplified OLS libraries

Assembly of plasmid library → Transformation of bacterial host → Selection and expansion → Analysis by sequencing

4 x 20 ul reactions:
75 ng vector
6.5 ng OLS library
4.5-fold molar excess of OLS library

95°C, 1 minute
95°C, 20 sec
60°C, 90 sec
65°C, 60 sec
8 cycles

23 minutes programmed
25-30 minutes real time

SPRI bead purification

Electroporation of
Electro Ten-blue cells
(20 electroporations)

Select in very low gelling agar

2-3 days at 30°C

MiSeq sequencing:
- **error rates**
- **distribution**

45 minutes,
15 minutes hands on

90 minutes
45 minutes hands on

2.5 days

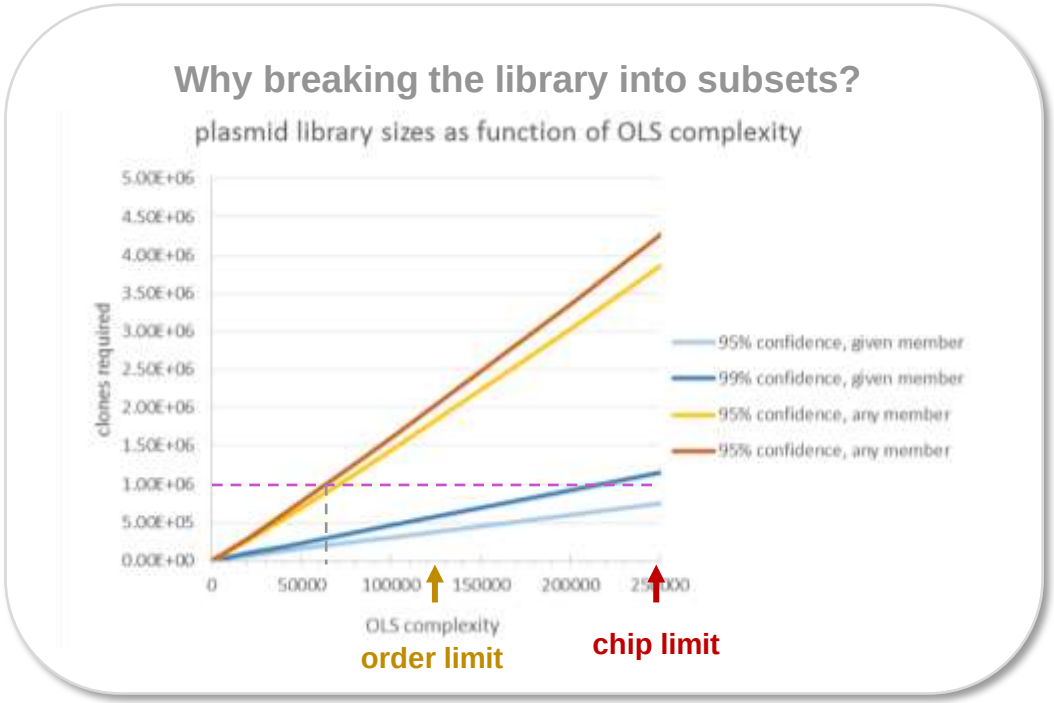
Evaluating the workflow: experimental set-up

Test case: GeCKo library for genome-wide CAS9 mediated gene knockout

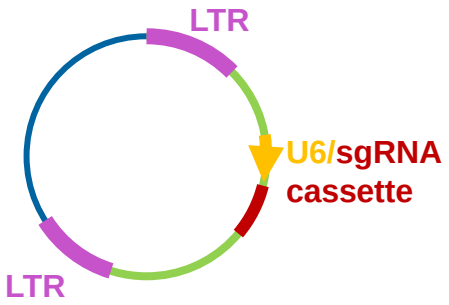
OLS library

GeCKo v2 libraries

species	human	mouse
genes targeted	19,050	20,611
targeting constructs/gene	6	6
miRNA targeted	1,864	1,175
targeting constructs/miRNA	4	4
control (nontargeting) sgRNAs	1,000	1,000
total sgRNA constructs	123,411	130,209
set A	66,172	
set B	57,239	
non-redundant set A	64,580	
non-redundant set B	56,869	
total non-redundant	121,449	



Recipient vector



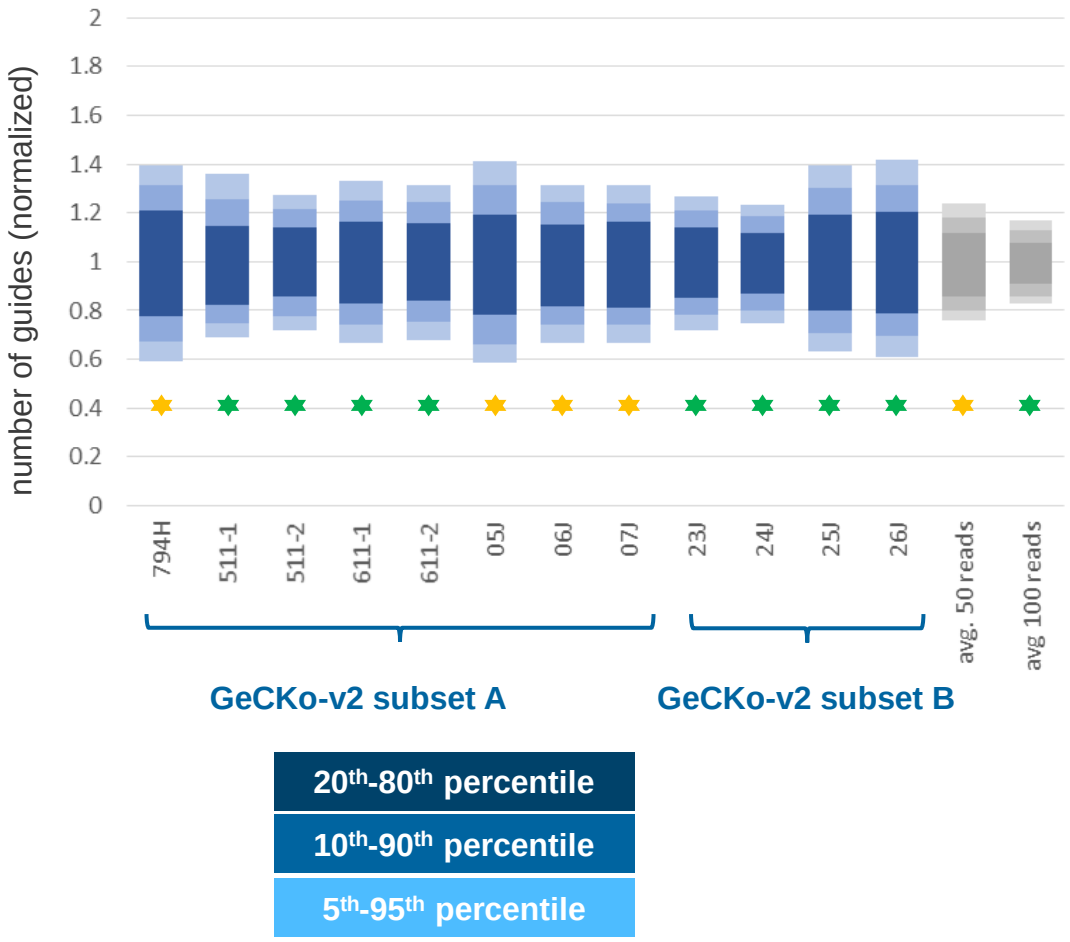
Retroviral vector
Size 6.5 kb
LTR(SIN) 226 bps

Lentiviral vector
Size 7.6 kb
LTR (SIN) 179 bps

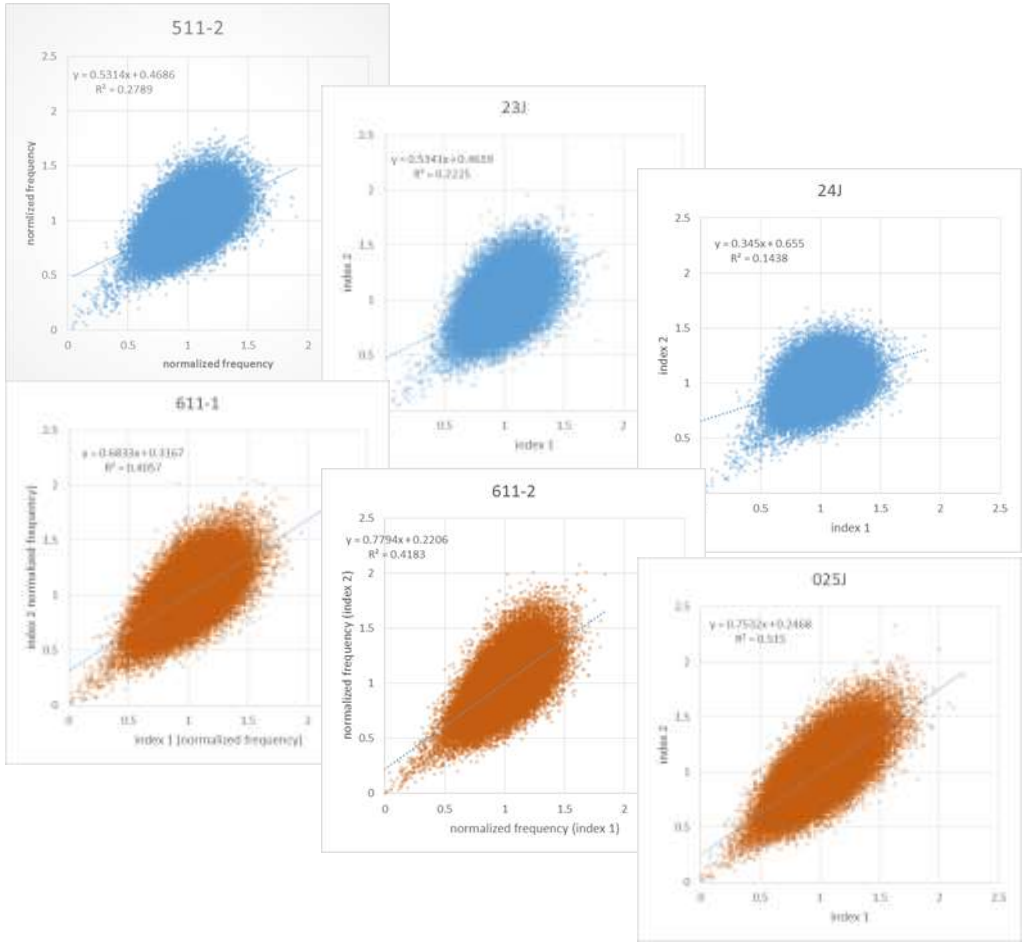
Sanjana et al., *Nature Methods* **11**: 783-784 (2014)

Quality: analysis of distribution in oligo libraries

Distribution of members analyzed using MiSeq

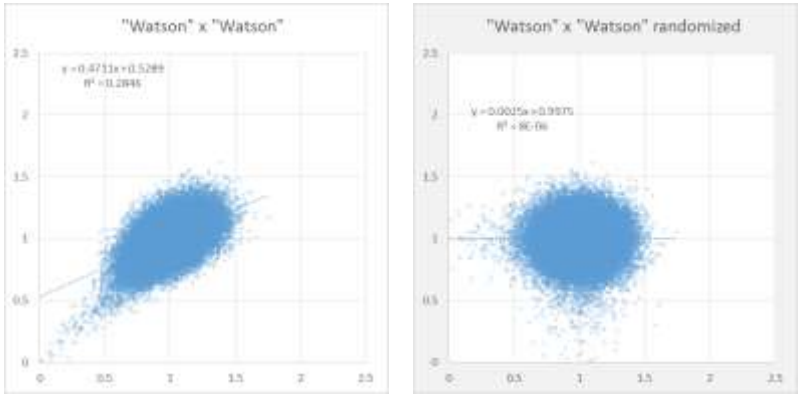


Typical bias in OLS libraries



Does strand polarity matter?

(+) strand library

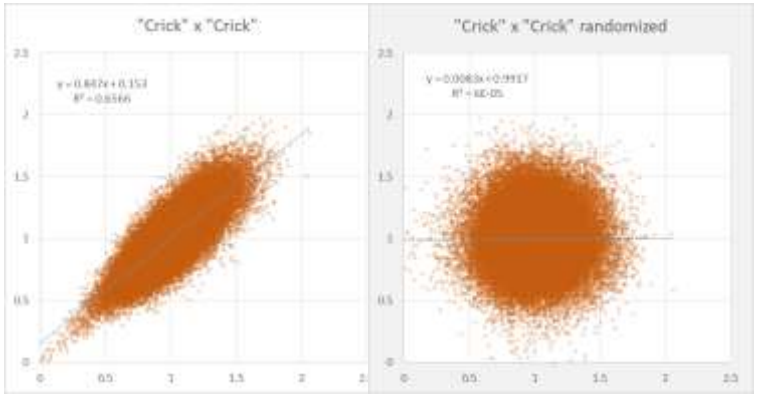


normalized by circular permutation

(+) strand OLS libraries and (-) strand OLS libraries show some systematic bias..

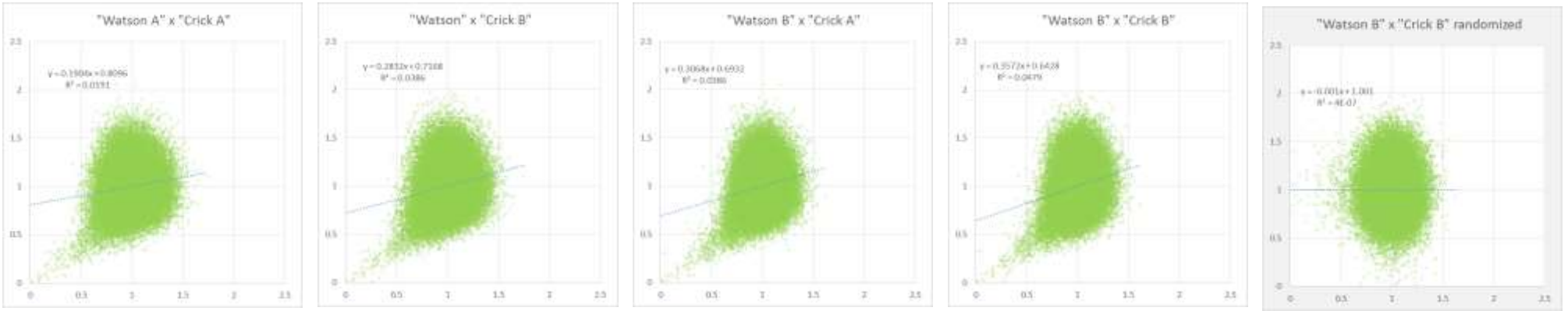
…but they are not similar to each other

(-) strand library



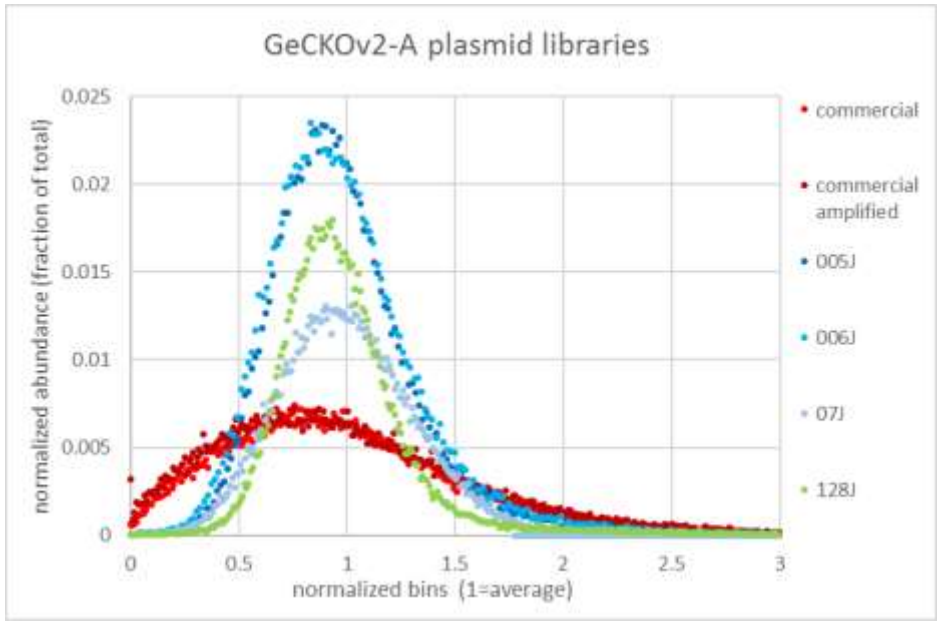
normalized by circular permutation

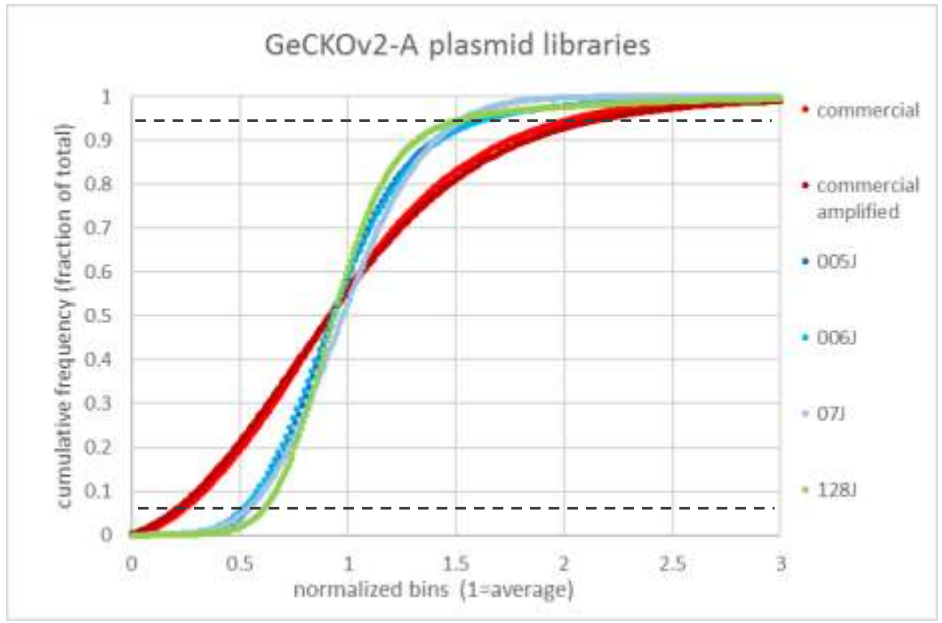
biases can be reduced by combining (+) and (-) strand libraries



normalized by circular permutation

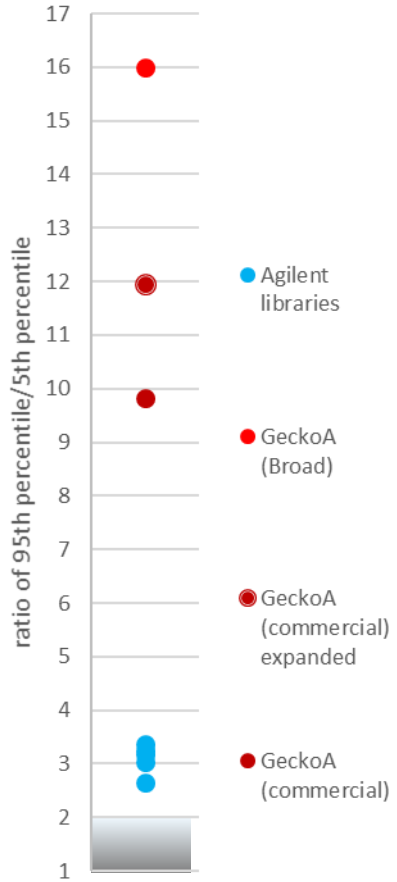
Quality of plasmid libraries: representation of library members





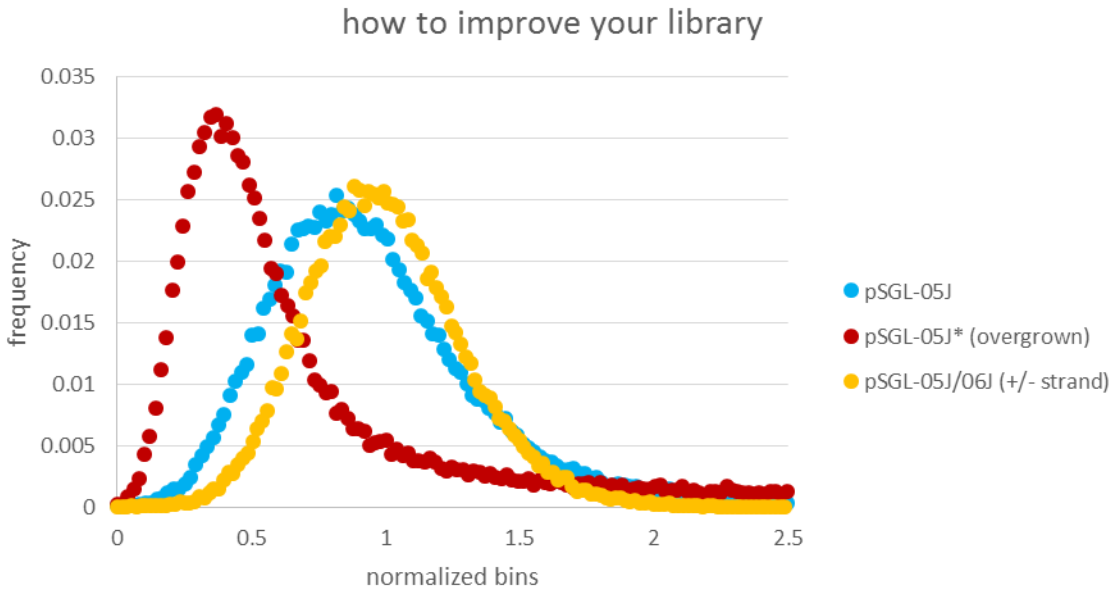
3 library constructions
3 different operators
3 different OLS

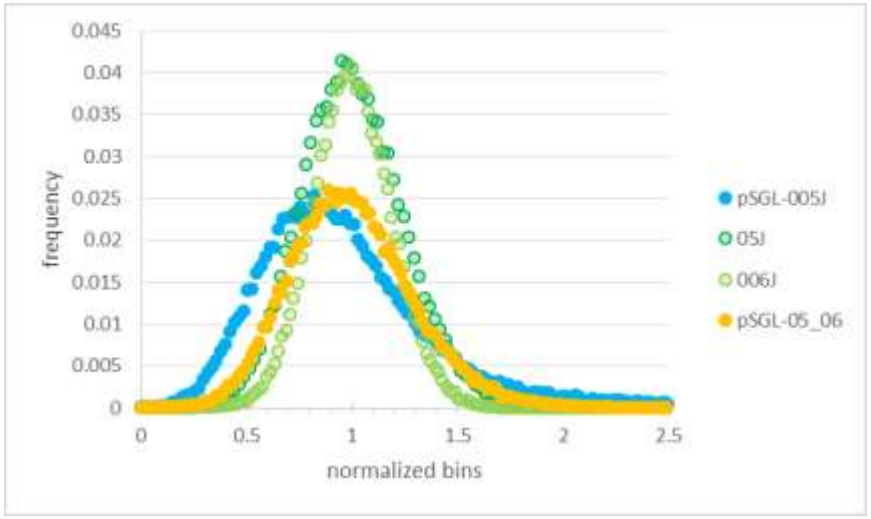
	missed guides	90th/10th percentile	95th/5th percentile	99.5th/0.5th percentile
pSGL-007J (SJ)	1	2.32	3.04	6.72
pSGL-006J (KF)	1	2.47	3.38	11.06
pSGL-005J (VZ)	1	2.38	3.19	9.83
pSGL-128J-dc (SJ)	1	1.99	2.64	8.30
GeckoA (Broad)	?	8.73	16.00	NA
GeckoA (commercial)	39	5.29	9.83	68.40
GeckoA (commercial) expanded	204	6.00	11.95	333.00



Generating highest quality plasmid libraries from OLS designs

- Start with highest quality OLS libraries
- combine (+) strand and (-) strand designs
- control growth conditions

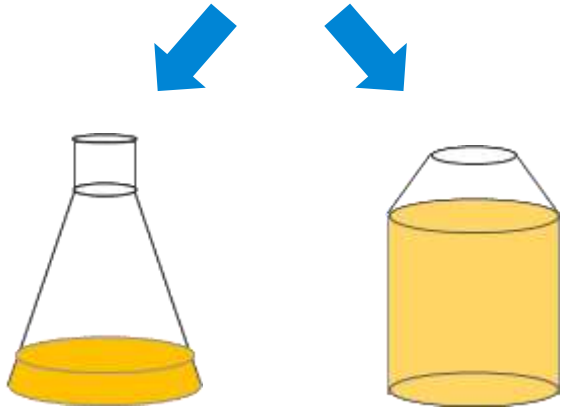




	pSGL-05J*	pSGL-05J	pSGL-05/06J
strand	(+)	(+)	(+/-)
99% range	NA	16.667	5.833
90% range	16.66666667	4.632	2.800
80% range	10.18181818	3.000	2.200
60% range	4	2.069	1.659
reads/guide	65.3	47.6	55.4
mode	0.398	0.819	0.884
median	0.536	0.903	0.974
average	1.000	1.000	1.000

Amplification of libraries degrades libraries by adding skew

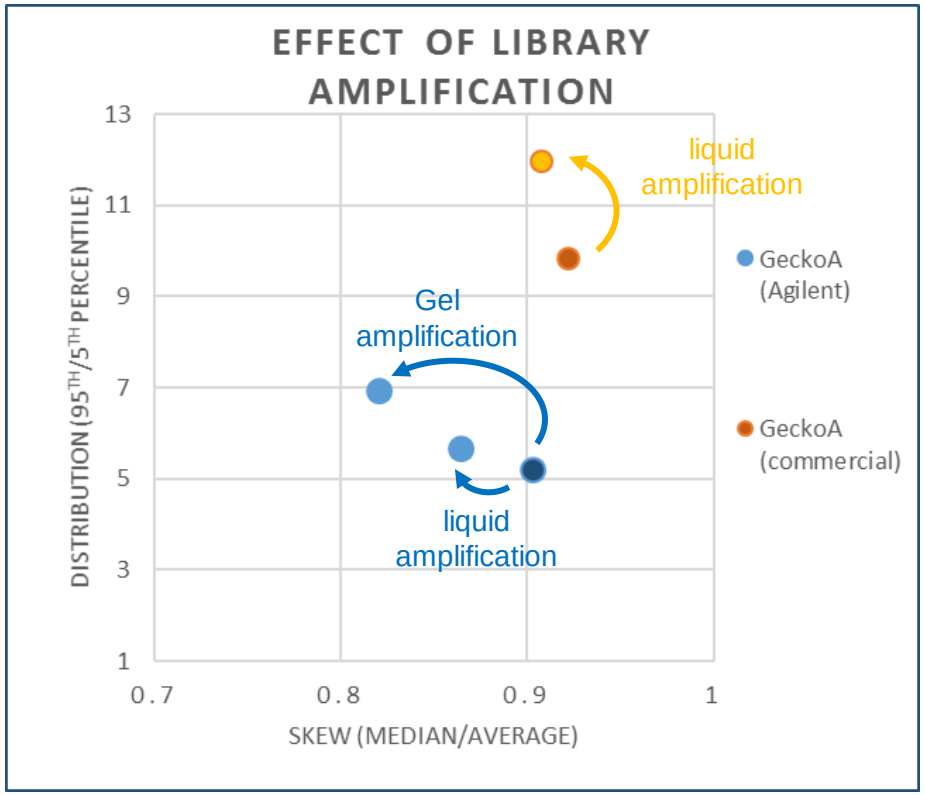
Transform
plasmid libraries



selection in liquid culture

selection in 3D matrix

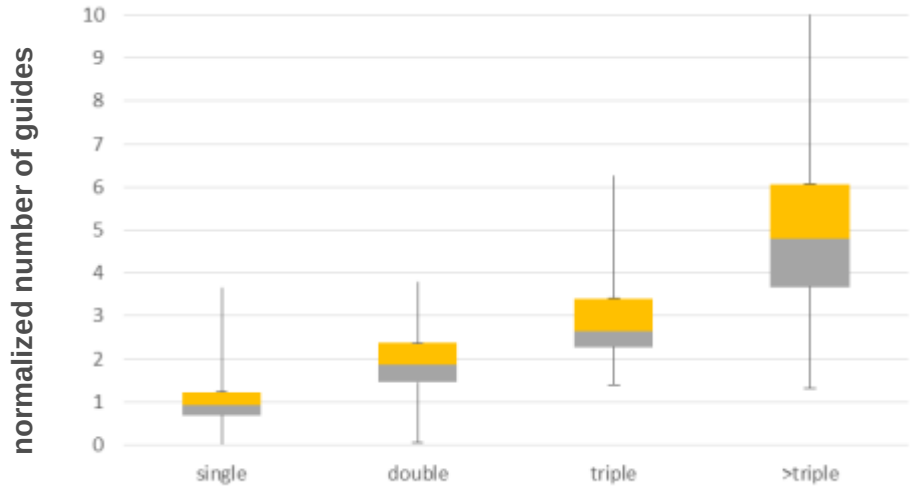
Analyze by NGS



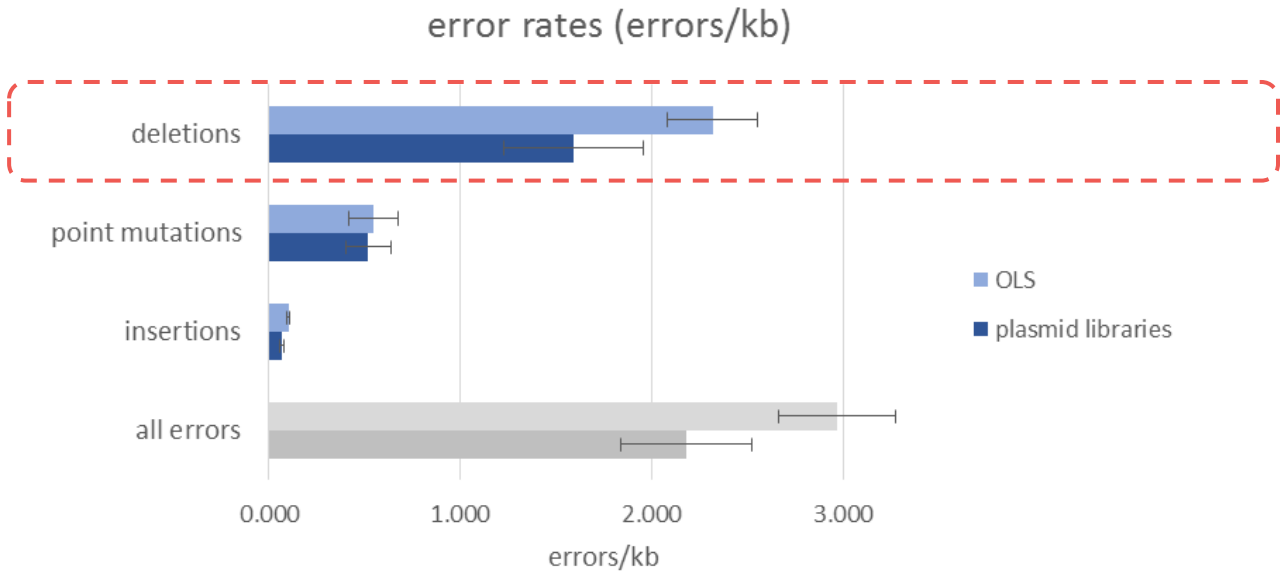
Designed biases are faithfully carried over into derived plasmid libraries

The original GeCKo v2 libraries contain redundant entries (2-22 fold)

- The relative ratios are maintained in the derived plasmid libraries
- The quality of the plasmid library depends on the quality of the underlying OLS
- Our library construction protocol can in principle be used to introduced designed biases through the OLS design



Quality II: fidelity of oligo library synthesis and derived plasmid libraries

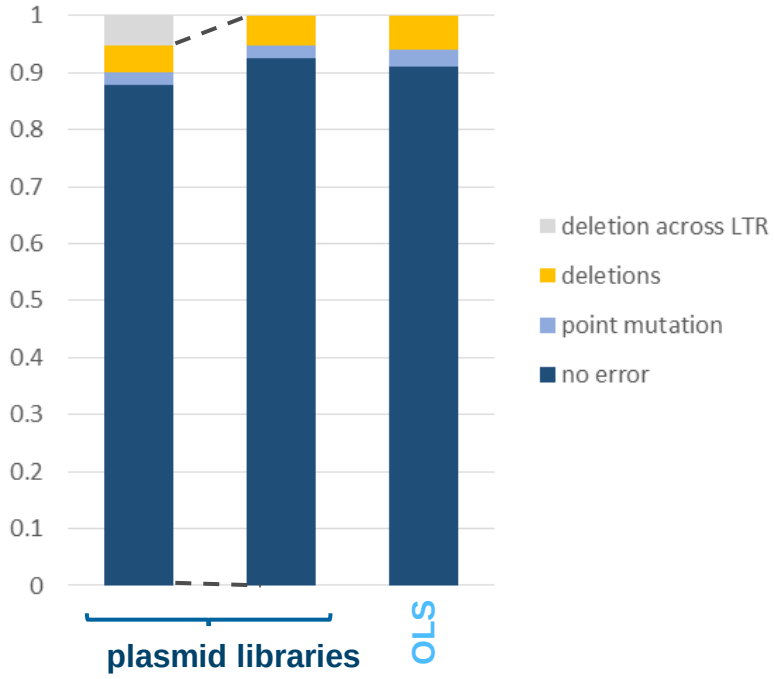


- Observed error rates are low (< 1 error/300 bases)
- predominant error is a deletion (expected)
- Point mutations are mostly resulting from the sequencing process (amplification and base calling confidence)
- Errors are biased against in the overlap assembly process (about 2/3 of OLS error rate)
- Error rate in the 20 nts library part is indistinguishable from the OLS error rate

	OLS libraries		plasmid libraries	
	error/kb	bases/error	error/kb	bases/error
deletions	2.32 ± 0.23	438 ± 73.5	2.18 ± 0.3	637 ± 84
point mutations	0.55 ± 0.13	1,870 ± 457	0.067 ± 0.007	2,022 ± 517
insertions	0.101 ± 0.007	10,064 ± 1403	0.52 ± 0.13	15,071 ± 1723
all errors	2.97 ± 0.3	311 ± 71.6	1.59 ± 0.23	485 ± 58

What matters in the end – what fraction of my library is correct?

- ≈ 90 of all clones contain no detectable error
- ≈ 5% of clones contain a deletion
- ≈ 2% of clones may contain a point mutation
- For retroviral/lentiviral vectors the recombination rate across the LTRs is ≈ 5%



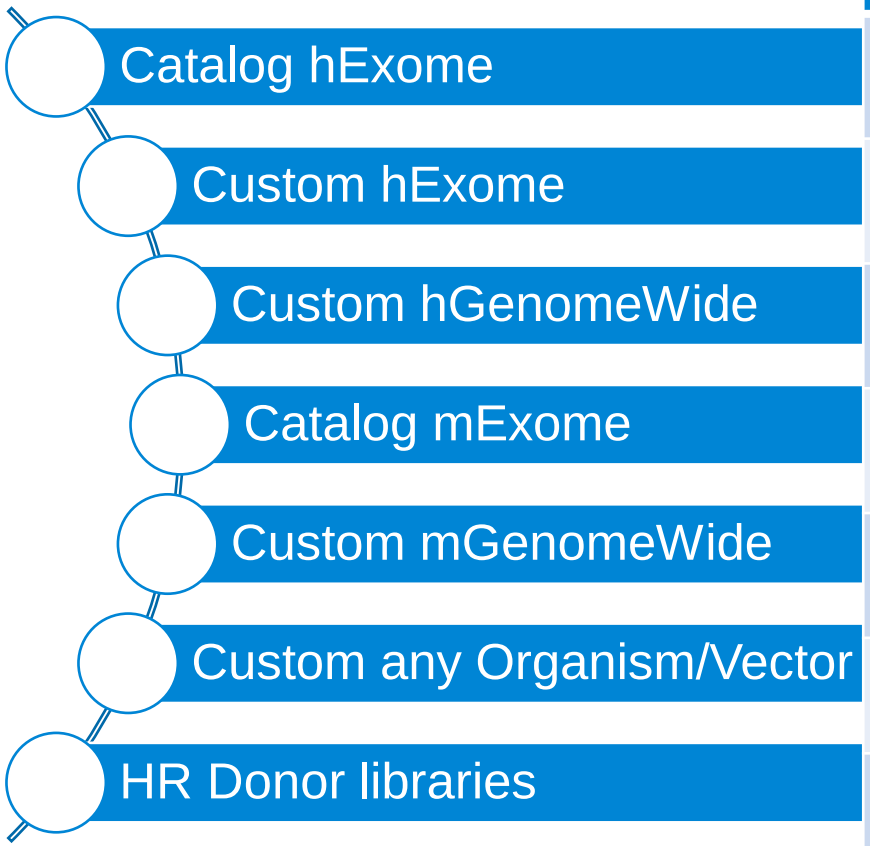
	plasmid libraries				OLS	
	average	stdev	average	stdev	average	stdev
n	4		4		10	
deletion across LTR	5.1%	2.3%			NA	NA
no error	87.8%	1.1%	92.6%	1.2%	91.1%	1.5%
point mutation	2.1%	0.4%	2.2%	0.4%	2.9%	1.1%
deletions	4.9%	0.8%	5.2%	0.9%	6.0%	1.3%

Conclusions

Agilent is a provider of high quality synthetic DNA oligomer libraries of complexities up to 128,000 different entries and lengths of up to 230 bases

High-efficiency cloning methods developed at Agilent allow fast and efficient construction of plasmid libraries based on OLS libraries with minimal additional bias.

SureGuide CRISPR libraries

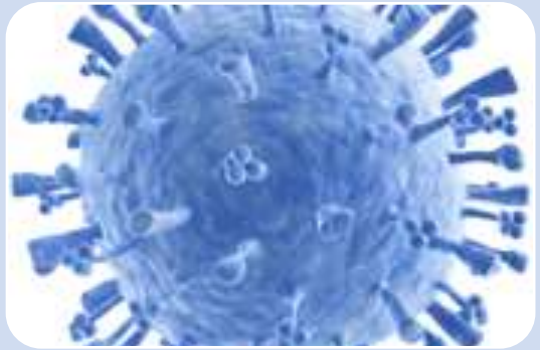


When (Early Access)	Usage	Application	Strengths
Now	Functional genomics/target ID	Knock-out libraries	Highest Quality, Good Value
Now	Functional genomics/target validation	Knock-out subsets	Highest Quality / Custom Designs
Now	**Genome regulatory networks**	CRISPR i/a	Highest Quality / Custom Designs
Now	Functional genomics/target ID	Knock-out libraries	Highest Quality / Custom Designs
Now	**Genome regulatory networks**	CRISPR i/a	Highest Quality / Custom Designs
Now	Various, fully custom	Customer design	Highest Quality / Custom Designs
Now	**Donor DNA for knock-ins, fully custom**	**Genome wide tagging, reporters**	**Quality/ Design freedom**

Enabling scientist to generate and test genomics hypothesis

SureGuide EA Products

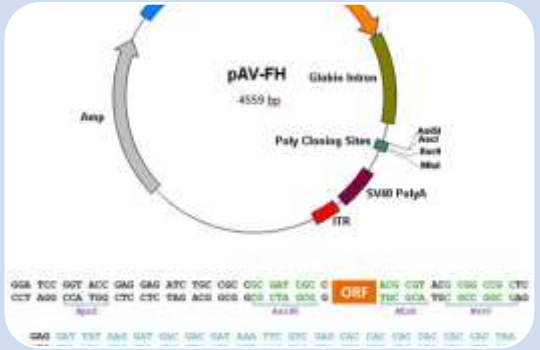
Ready-to-Package



Catalog Libraries

- Plasmid Library
- GeCKOv2
- Human and Mouse
- Cloned into lentivirus vector with hU6 promoter

Ready-to-Clone



Custom Libraries

- Pre-amplified OLS library
- User defined subset or designed

Human and mouse

- Compatible with SureVector cloning

Ready-to-Amplify

Custom Libraries

- Unamplified oligo pool
- Any species, any cloning method
- Entirely custom by user design

Acknowledgments

La Jolla
Katie Felts
Sarah Johns
Vivian Zhang
Peter Sheffield

Santa Clara
Magnus Isaksson
Ben Borgo