Oligo Pools:

Design, Synthesis, and Research Applications

Presenter：Marcelo Caraballo, Senior Scientist of CustomArray

Date　　：December 13, 2018

CONTENTS

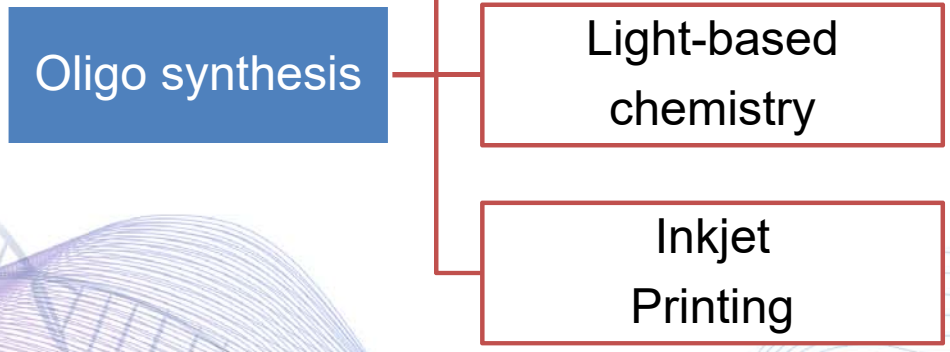
CONTENTS

What is an oligo pool?

- Using an **in situ array** technology, the synthesis of oligonucleotides can benefit from the same parallelization that has revolutionized the DNA sequencing field.

- The end product is a **library** of thousands to hundreds of thousands oligos that is completely defined by the customer at a tiny fraction of the cost of making each oligo individually via traditional oligo synthesis techniques.

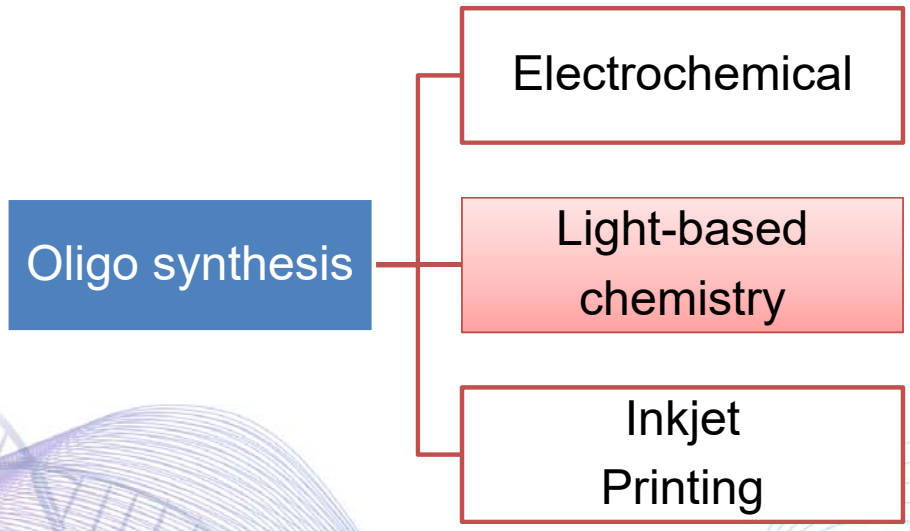
In situ array synthesis



All in situ array oligo synthesis technologies have to solve the same central question.

How to spatially segregate chemical reactions on a planar, or mostly planar, surface without using physical containment (walls)?

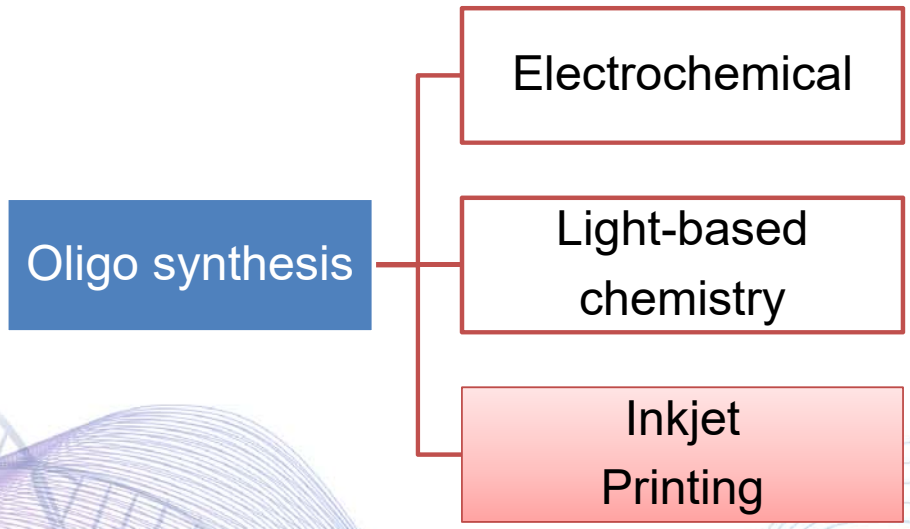
The 3 main branches of array synthesis



Light-based array synthesis

- Uses custom phosphoramidites with light sensitive protecting groups (NPOC) and localized light (photolithography, DLP, laser light, etc) to perform the spatial segregation.
- Expensive and poor synthesis fidelity. Some incarnations suffered from high equipment costs.
- Largely abandoned now.

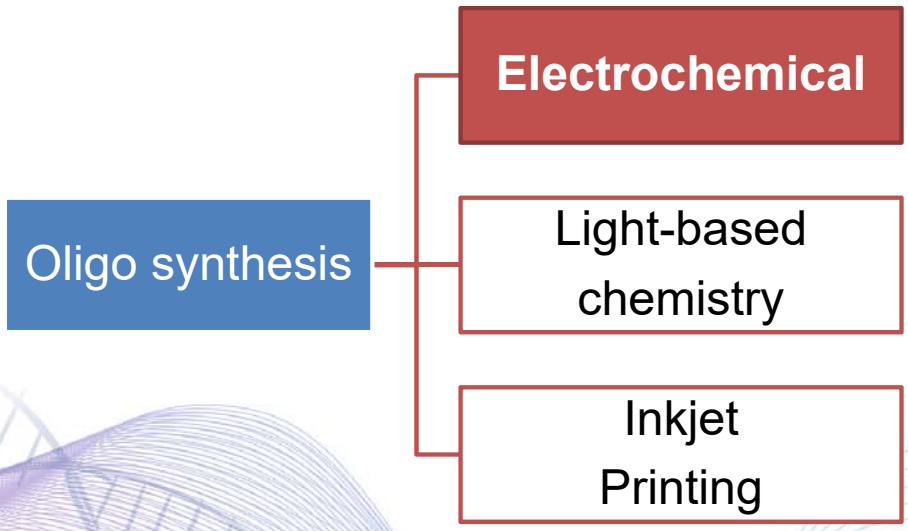
The 3 main branches of array synthesis



Inkjet printing

- Benefits from off-the-shelf reagents that are inexpensive and very reproducible
- Suffers from workflow bottlenecks due to equipment restrictions
- Large complicated, high-tech inkjet printing devices are difficult to build, maintain, and operate

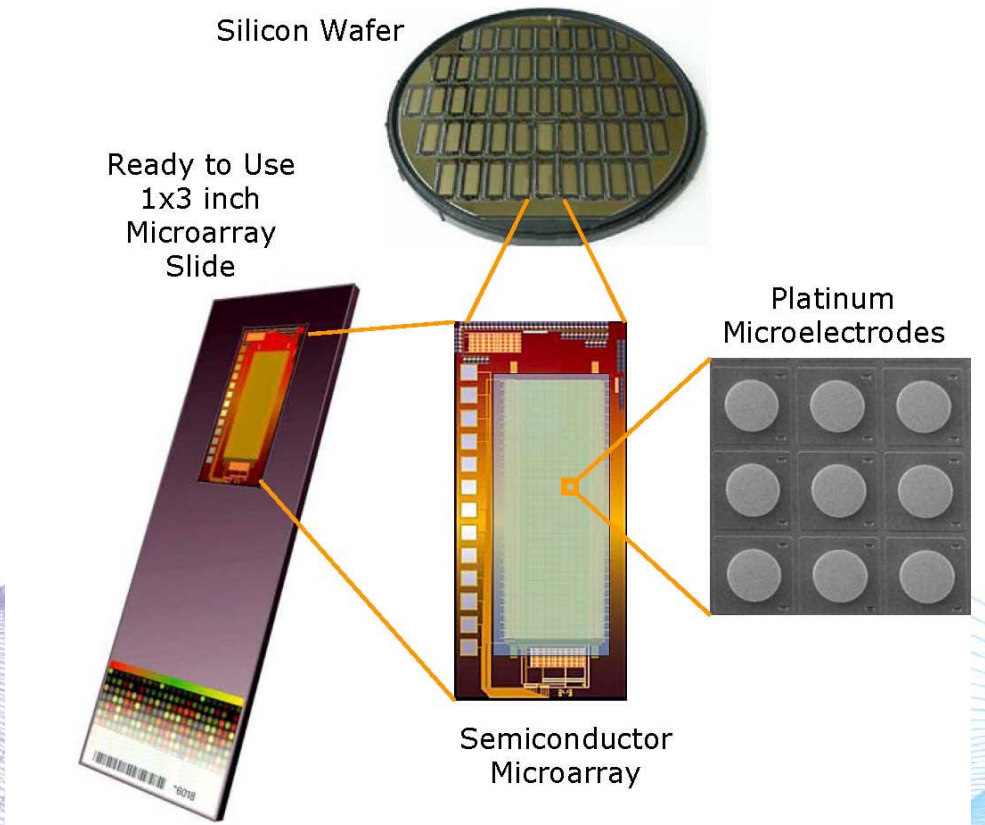
The 3 main branches of array synthesis



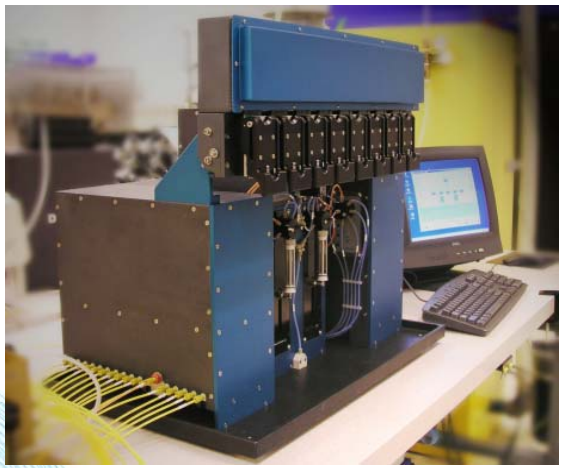
Electrochemical synthesis

- Leverages the semi-conductor industry to achieve the most reproducible and high-throughput synthesis possible
- The “chip” is the technology, whereas the synthesizer is a simple fluid mover
- Chips can be made by tens of thousands easily by any semi-conductor foundry
- Fastest synthesis in the industry with high sequence fidelity due to the flexibility of using a simple synthesizer with an advanced chip

Electrochemical oligo synthesis using CMOS technology



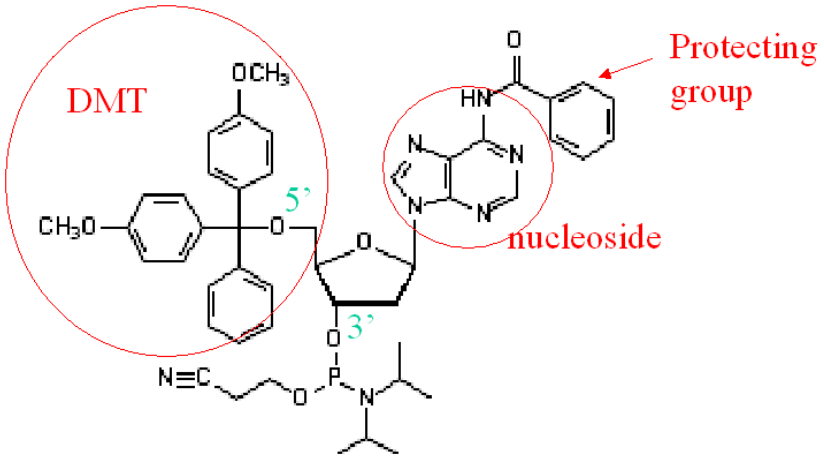
- ✓ Software applies voltage to sets of specific electrodes
- ✓ Electrode activation controls chemical reactions at each individual electrode on the microarray



But first, a primer on oligo synthesis

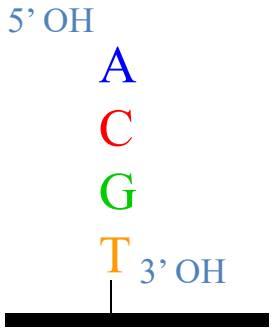
- The current incarnation of chemical oligo synthesis dates back to Marvin Caruthers at the University of Colorado, Boulder in the early 1980's.

- Various modifications and improvements have followed, but all current chemical oligo synthesis processes flow directly from that landmark work.



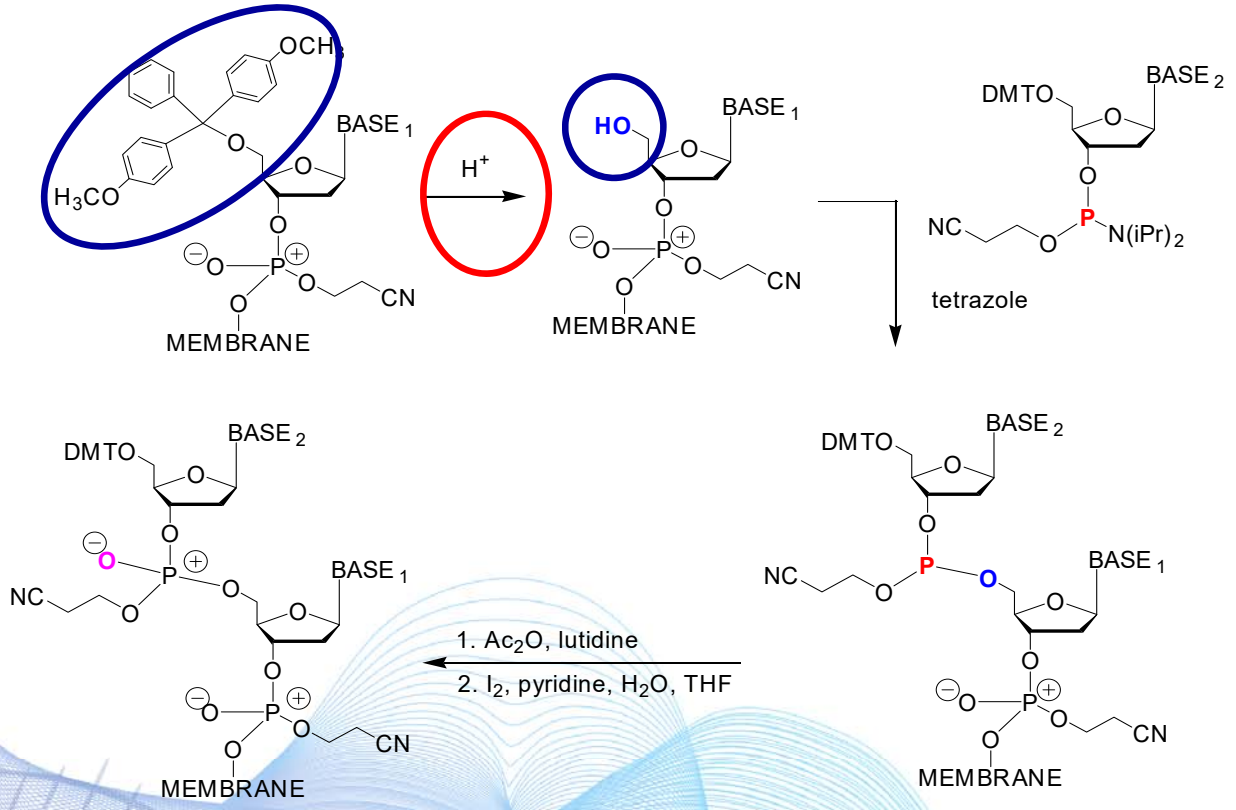
4 steps to add one nucleotide

- Each nucleotide addition requires 4 steps
  - Detritylation
  - Activation and Coupling
  - Capping
  - Oxidation



- Repeat steps for next nucleotide

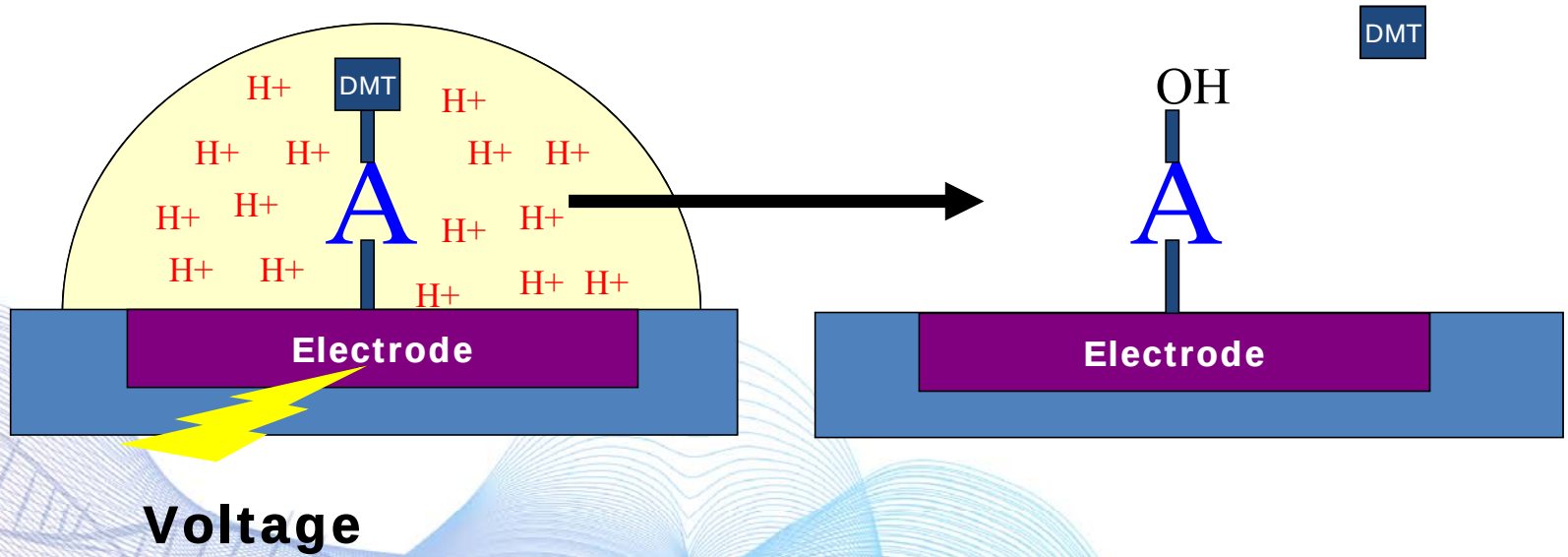
Normal phosphoramidite chemistry with electrochemical deprotection



BASE = protected A or T or G or C

CustomArray™ technology

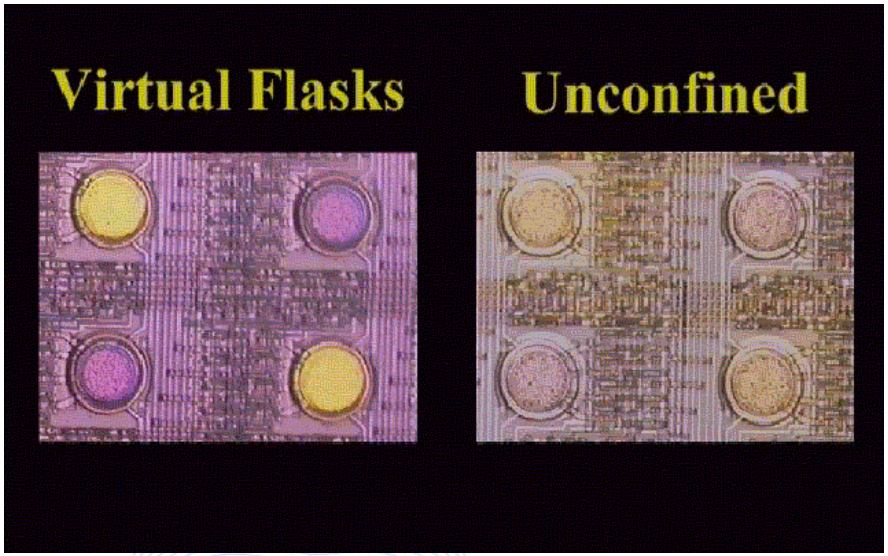
- Detritylation requires acid (H+), TCA in MeCl2
- CustomArray generates acid electrochemically at the electrode surface



Proton confinement

Minimize H^+ half-life distance

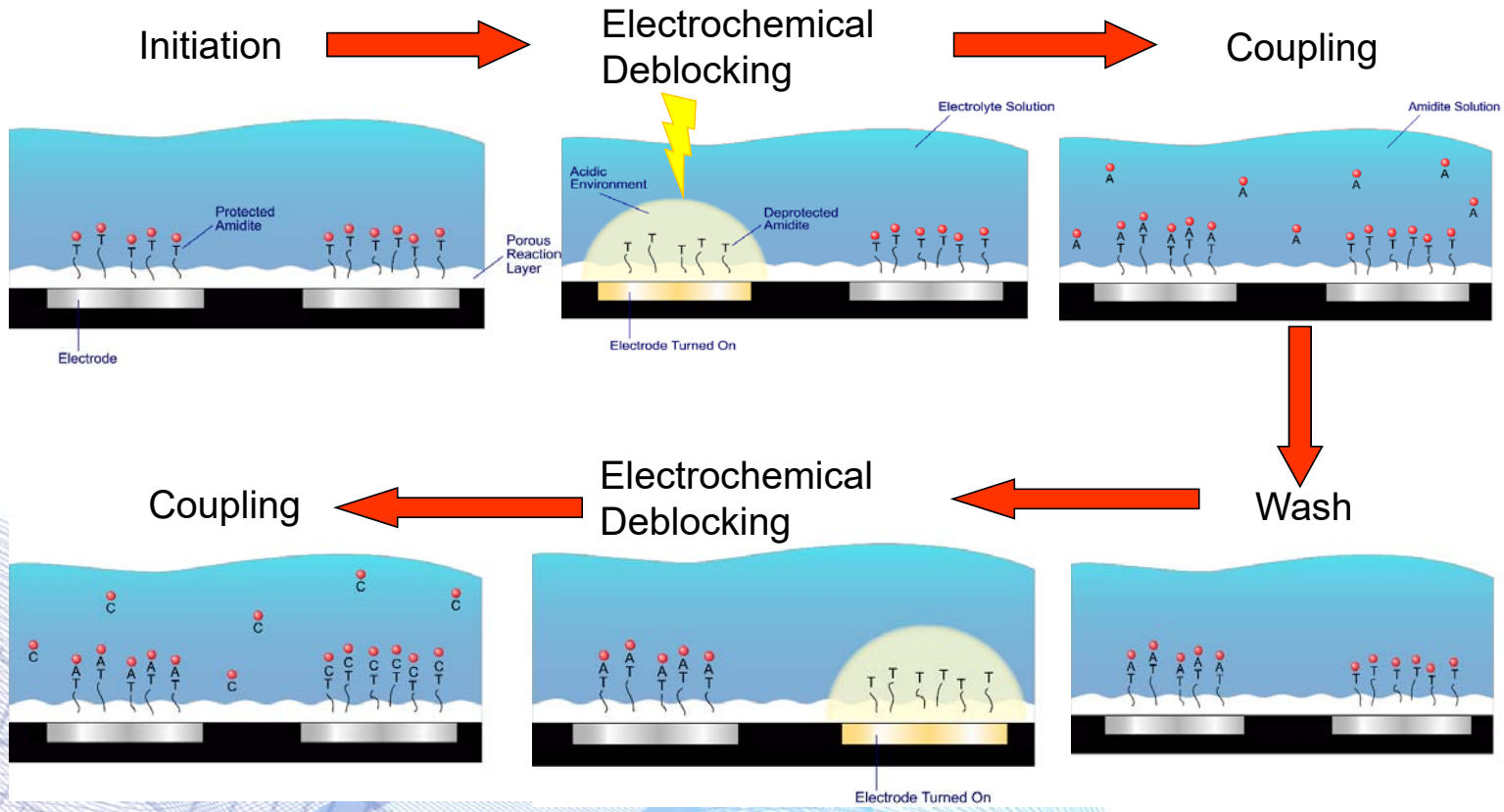
Acid confined above electrode



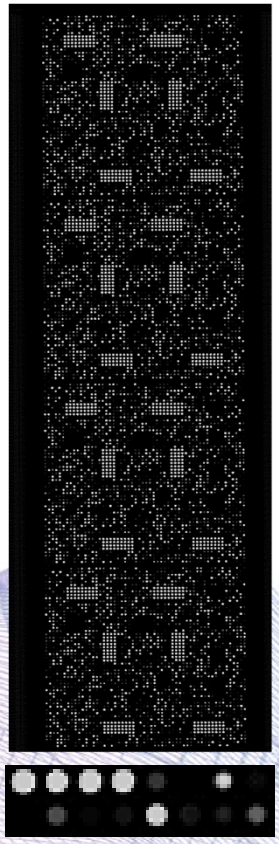
Acid diffused away from electrode

Bromophenol Blue dye added for illustrative purposes

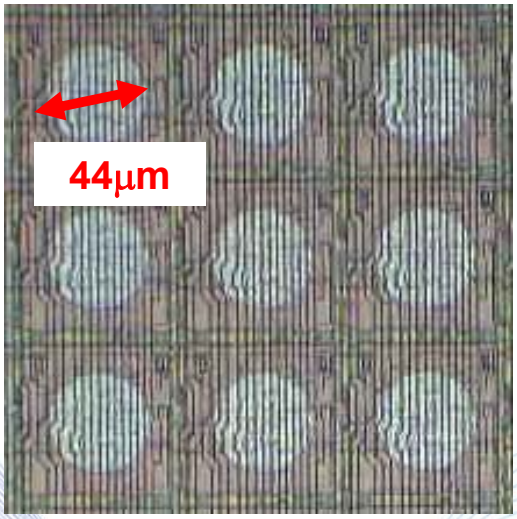
Electrochemical synthesis

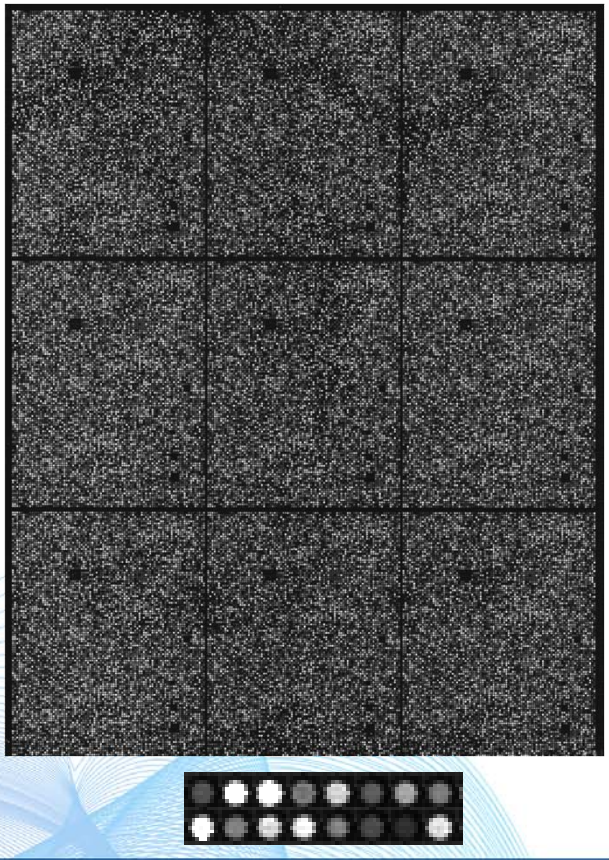


CustomArray™ versions

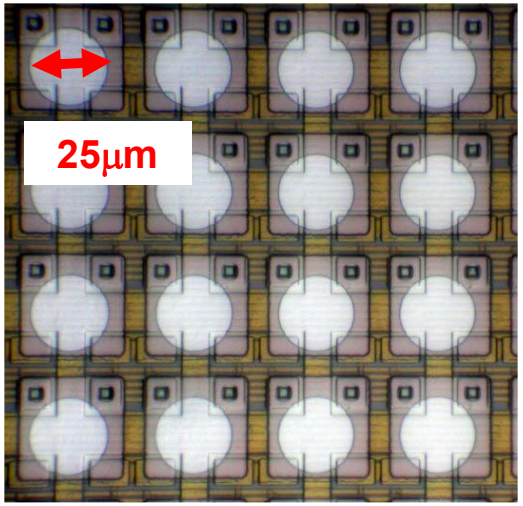


12K CustomArray
Up to 12,472 oligos



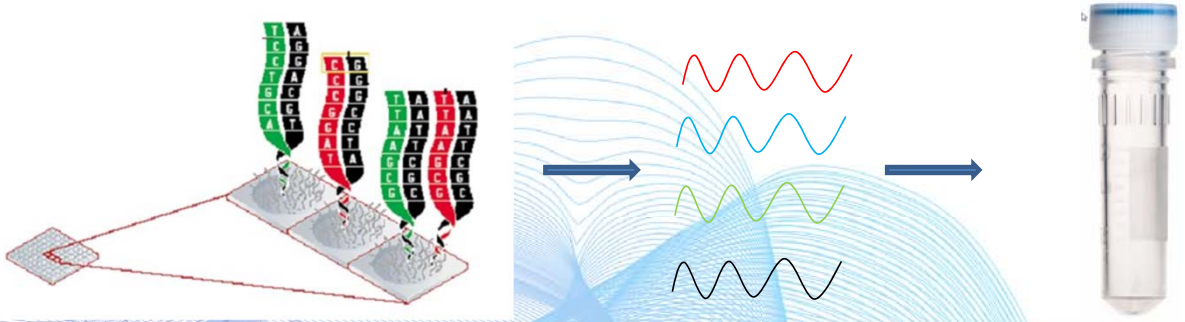


90K CustomArray
Up to 92,918 oligos



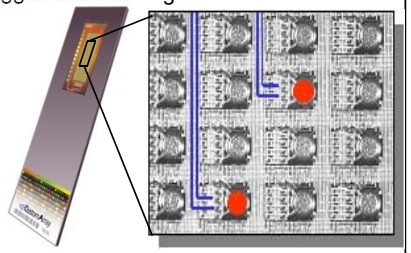
How are oligo pools constructed?

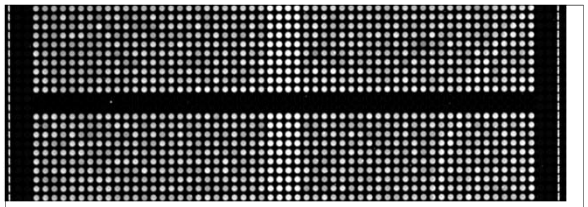
1. Using the electrode array as a starting point, we individually synthesize oligonucleotides at each point of the array
2. After the synthesis is complete oligos are removed from the surface into a common tube
3. After some minor processing, we ship the oligos as ssDNA suspended in TE buffer



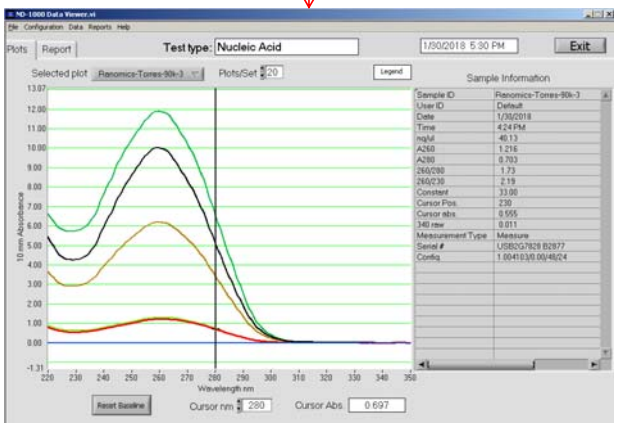
Quality control of oligo pools

- Semi-Conductor technology allows for electronic verification of electrode activation.
- Performance of all electrodes is verified and logged for each oligo pool.



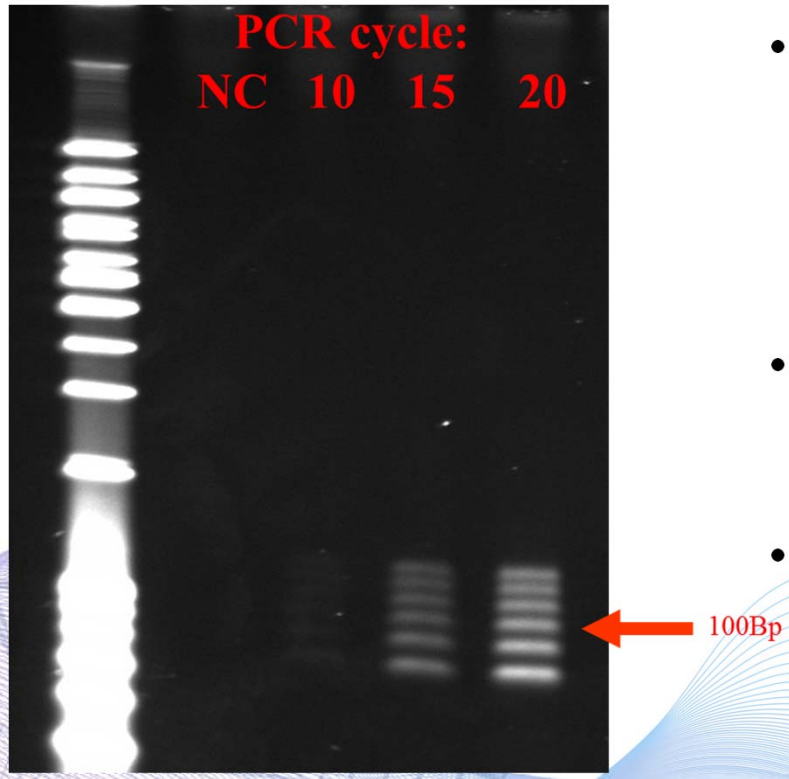


Post synthesis, presence of DNA can be visually verified for each electrode. Blank electrodes are intentional to provide contrast.



Additional QC can be done via PCR. This is done on test pools that run alongside customer orders.

Quality control of oligo pools



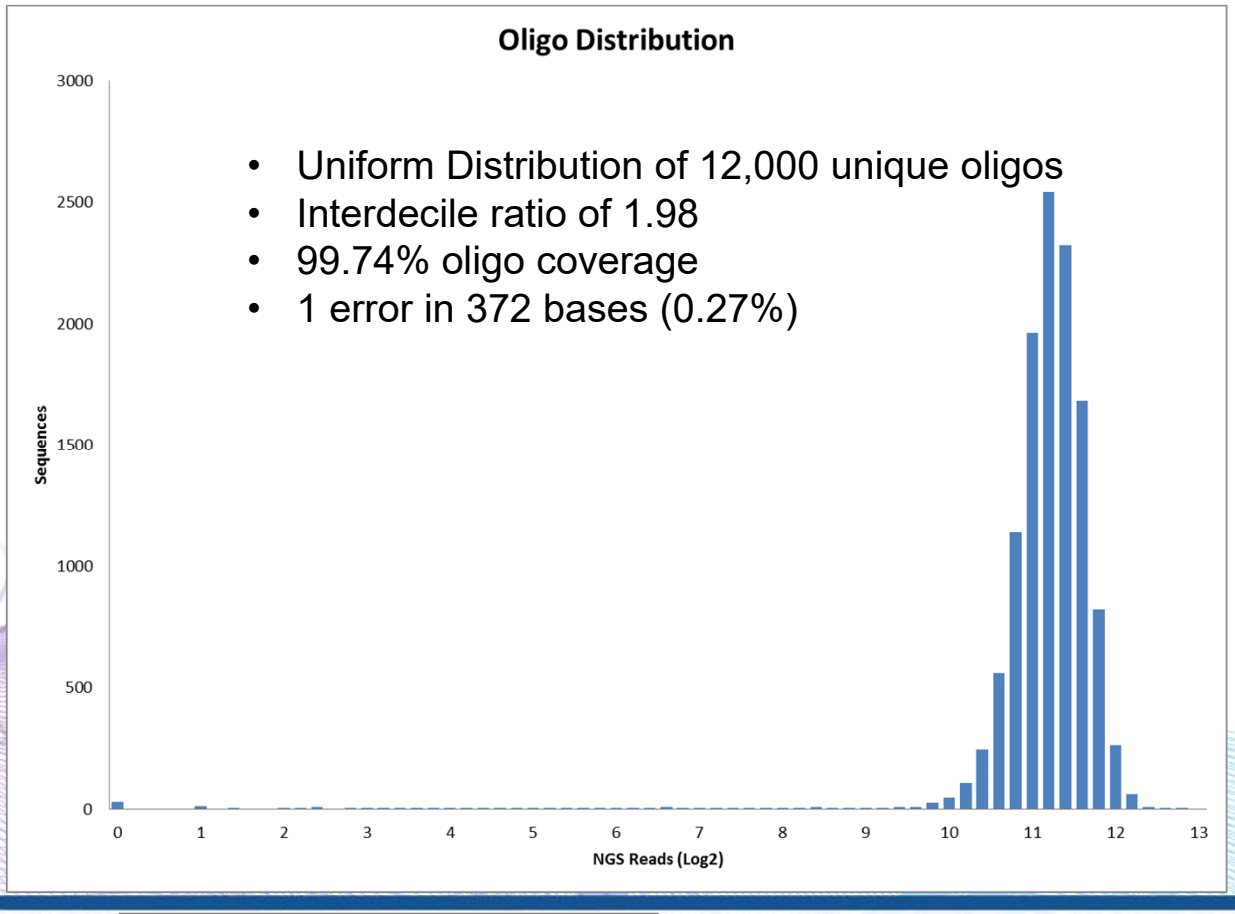
- Oligo pool product amplification via universal PCR confirms the presence of 80, 90, 100, 110, 120, and 130 bp oligonucleotides synthesized on one 12k Microarray.
- Any set of sequences can be written on the chip as long as amplification primers are included.
- Once a pool has been generated, a large supply of oligos can be subsequently generated by PCR amplification.

Oligo pool stability study

- 110mer, 12K Oligo Pools – same file
- Spanned 8 weeks using different reagents, machines, chip lots, etc.
- Data provided by customer using NGS
- Error rates range from 0.43% to 0.73%

chip	% perfect	22mer % perfect	70% fold Difference	90% fold difference	% Recovery
1	65.8	91.3	1.7	2.6	99.3
2	70.8	92.7	1.6	2.3	99.3
3	63.4	91.2	1.7	2.7	99.3
4	65.9	90.0	1.4	2.0	100.0
5	58.2	89.1	1.4	2.0	99.3
6	66.7	92.3	1.5	2.2	99.3
7	64.0	91.2	1.5	2.0	99.3
8	55.6	89.3	1.5	2.1	99.3

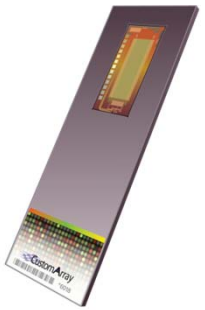
Distribution case study



- Uniform Distribution of 12,000 unique oligos
- Interdecile ratio of 1.98
- 99.74% oligo coverage
- 1 error in 372 bases (0.27%)

Case Study of Customer Submitted Sequences

- Customer ordered 12,000 125mer oligo pool from CustomArray
- Using our proprietary CMOS semi-conductor array synthesis platform, we performed the synthesis and delivered the product within 7 days
- Post synthesis, we minimally amplified the oligo pool and performed NGS using an Illumina Hi-Seq



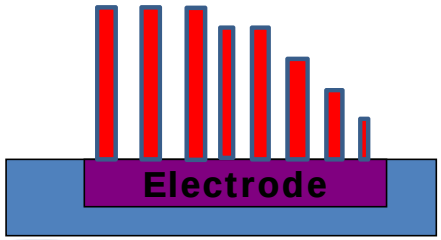
CONTENTS

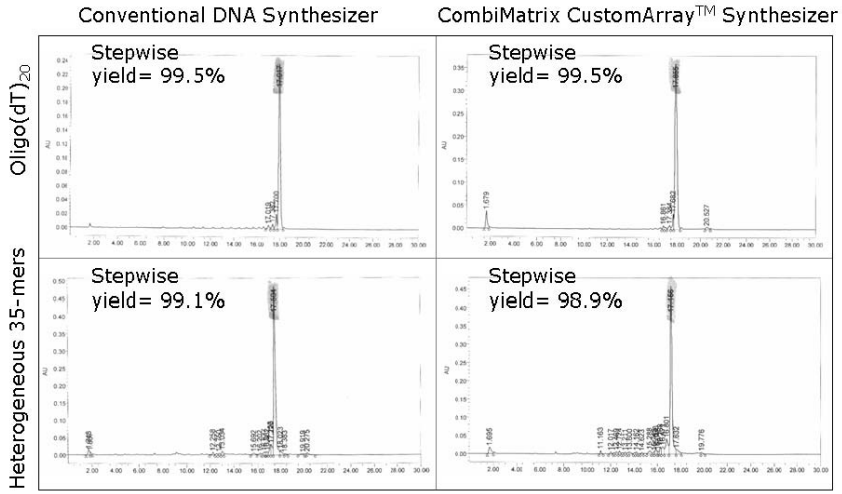
First things first, amplification

Oligo pools generally need to be amplified as a first step

There are 3 main reasons:

- Purification
- Low initial copy numbers
- QC





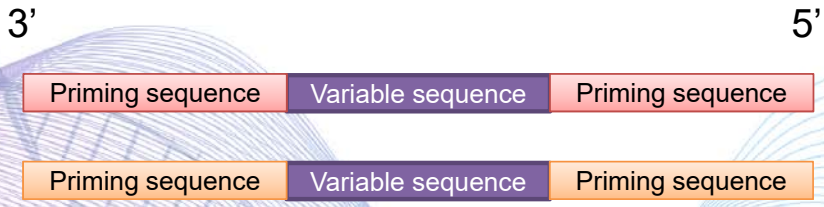
All oligo synthesis suffers from a step-wise truncation every time a base is added.

Probe Length	% of Full-Length Oligonucleotides
20-mer	81
40-mer	70
60-mer	50

Proportion of full-length oligonucleotides among probes of different length synthesized with 99% step-wise efficiency.

PCR amplification

- 1st cyle of PCR converts ssDNA into dsDNA.
- Subsequent cycles only amplify full-length copies that include 5’ priming sequence

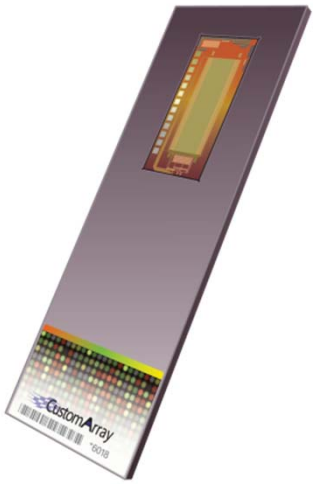


Priming sequences can be universal to allow all oligos to be amplified by a single pair of primers or a pool can be subdivided into an arbitrary number of sub-pools, each with an arbitrary number of oligos, by using different priming sequences.

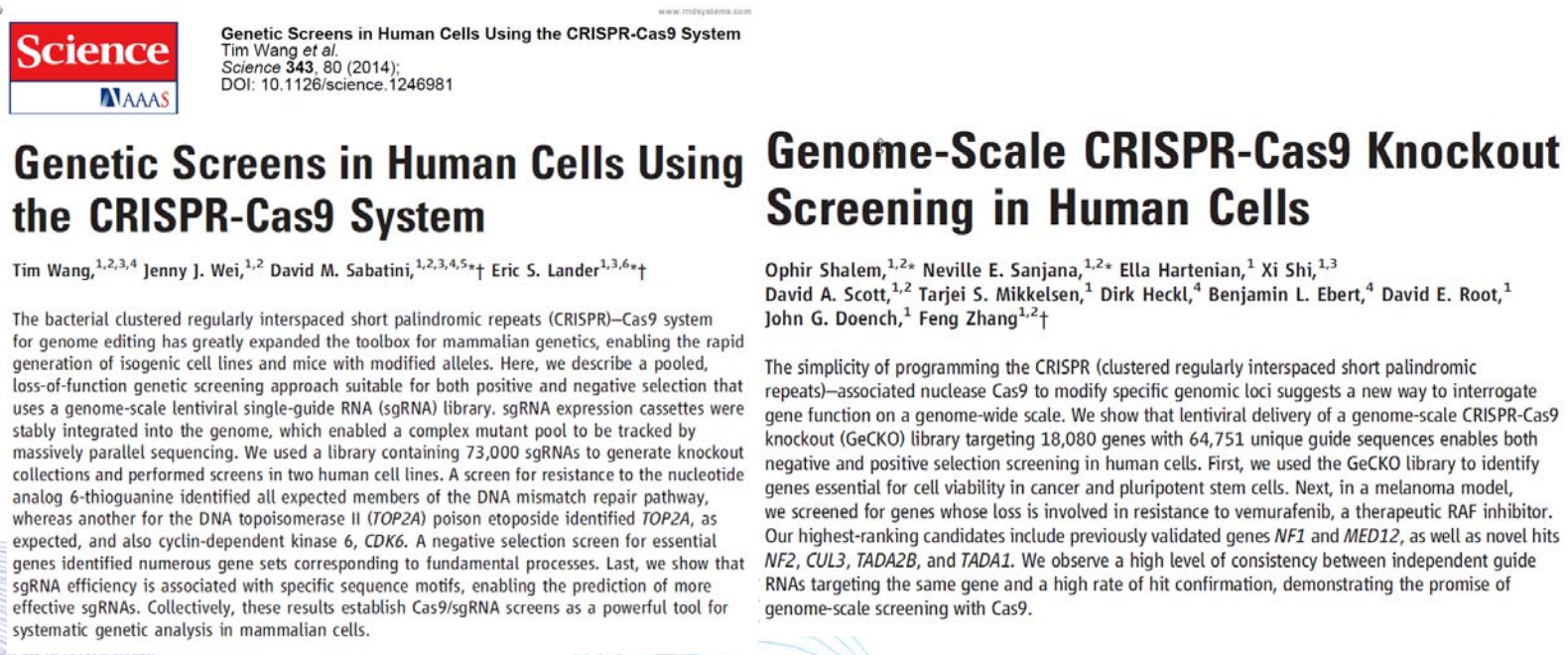
CONTENTS

What are oligo pools used for?

- Genome editing libraries
  - ✓ **CRISPR gRNA screening libraries**
  - ✓ shRNA screening libraries
- **Targeted sequencing**
  - ✓ Hybrid-capture
  - ✓ MIP style
- **Mutagenesis libraries**
- OligoFish (Merfish)
  - ✓ In situ hybridization applications
- **DNA data storage**
- MPRA (Massively Parallel Reporter Assay)



CRISPR gRNA libraries



Science
AAAS

Genetic Screens in Human Cells Using the CRISPR-Cas9 System
Tim Wang *et al.*
Science **343**, 80 (2014);
DOI: 10.1126/science.1246981

Genetic Screens in Human Cells Using the CRISPR-Cas9 System

Tim Wang,[1,2,3,4] Jenny J. Wei,[1,2] David M. Sabatini,[1,2,3,4,5]*† Eric S. Lander[1,3,6]*†

The bacterial clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 system for genome editing has greatly expanded the toolbox for mammalian genetics, enabling the rapid generation of isogenic cell lines and mice with modified alleles. Here, we describe a pooled, loss-of-function genetic screening approach suitable for both positive and negative selection that uses a genome-scale lentiviral single-guide RNA (sgRNA) library. sgRNA expression cassettes were stably integrated into the genome, which enabled a complex mutant pool to be tracked by massively parallel sequencing. We used a library containing 73,000 sgRNAs to generate knockout collections and performed screens in two human cell lines. A screen for resistance to the nucleotide analog 6-thioguanine identified all expected members of the DNA mismatch repair pathway, whereas another for the DNA topoisomerase II (*TOP2A*) poison etoposide identified *TOP2A*, as expected, and also cyclin-dependent kinase 6, *CDK6*. A negative selection screen for essential genes identified numerous gene sets corresponding to fundamental processes. Last, we show that sgRNA efficiency is associated with specific sequence motifs, enabling the prediction of more effective sgRNAs. Collectively, these results establish Cas9/sgRNA screens as a powerful tool for systematic genetic analysis in mammalian cells.

Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells

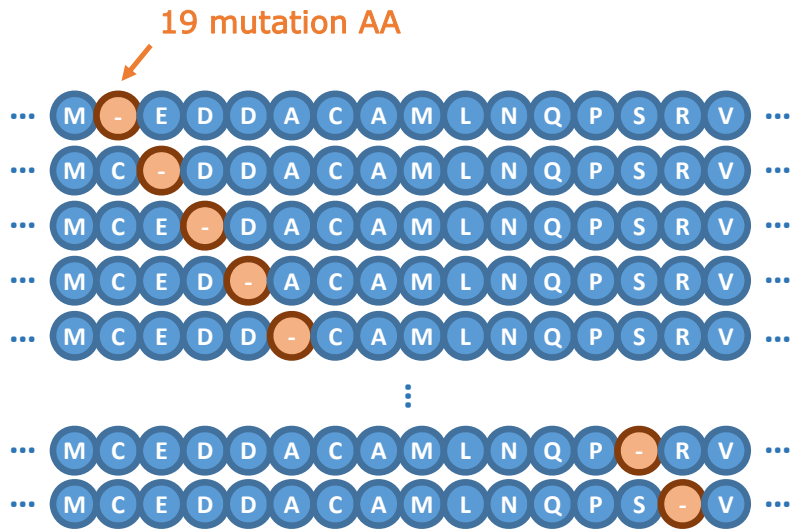
Ophir Shalem,[1,2]* Neville E. Sanjana,[1,2]* Ella Hartenian,[1] Xi Shi,[1,3] David A. Scott,[1,2] Tarjei S. Mikkelsen,[1] Dirk Heckl,[4] Benjamin L. Ebert,[4] David E. Root,[1] John G. Doench,[1] Feng Zhang[1,2]†

The simplicity of programming the CRISPR (clustered regularly interspaced short palindromic repeats)–associated nuclease Cas9 to modify specific genomic loci suggests a new way to interrogate gene function on a genome-wide scale. We show that lentiviral delivery of a genome-scale CRISPR-Cas9 knockout (GeCKO) library targeting 18,080 genes with 64,751 unique guide sequences enables both negative and positive selection screening in human cells. First, we used the GeCKO library to identify genes essential for cell viability in cancer and pluripotent stem cells. Next, in a melanoma model, we screened for genes whose loss is involved in resistance to vemurafenib, a therapeutic RAF inhibitor. Our highest-ranking candidates include previously validated genes *NF1* and *MED12*, as well as novel hits *NF2*, *CUL3*, *TADA2B*, and *TADA1*. We observe a high level of consistency between independent guide RNAs targeting the same gene and a high rate of hit confirmation, demonstrating the promise of genome-scale screening with Cas9.

Many available CRISPR libraries were derived from oligos originally made using CustomArray technology

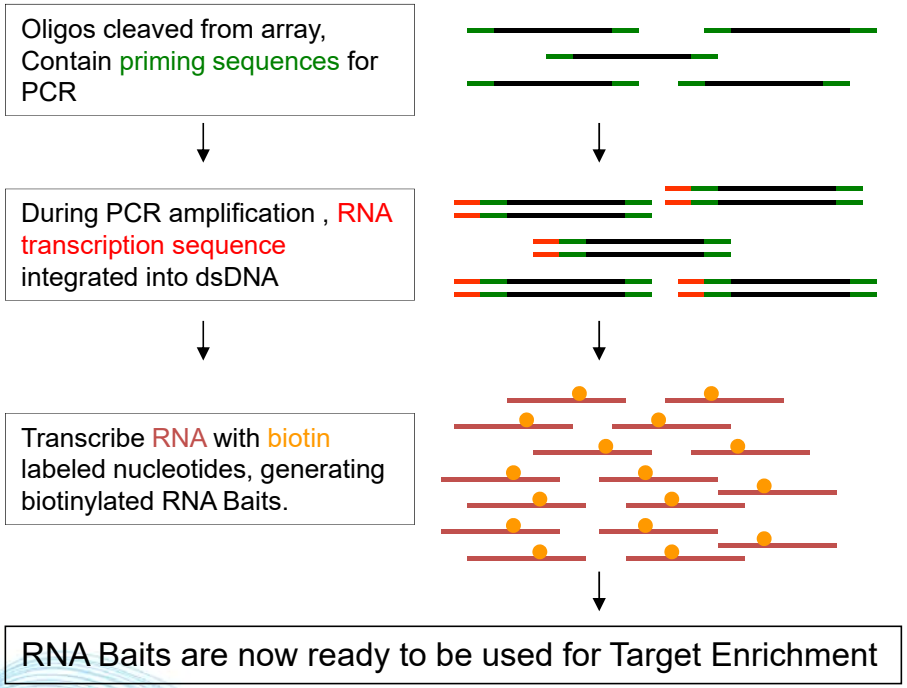
Gene variant libraries

- Synthesize thousands of variants at one time to identify structural and functional residues and to optimize protein function
- Ideal for protein engineering, industrial enzyme development, and metabolic engineering
- **Types of gene variant libraries:**
  - Site-directed mutagenesis
  - Site-saturation mutagenesis
  - Saturation scanning mutagenesis
  - Combinatorial mutagenesis



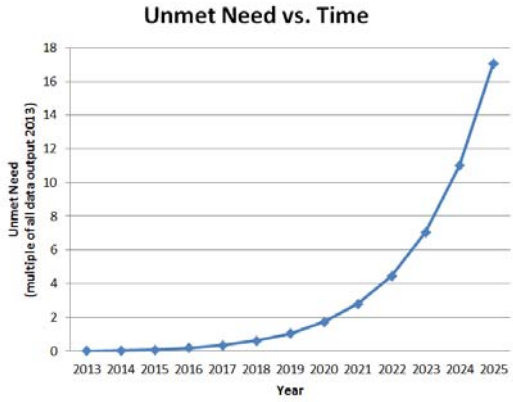
Target enrichment: bait and capture

- Capture the targeted sequence using Biotinylated-RNA or DNA bait molecules in solution instead of DNA fixed to a surface.
- **Benefits**
  - Very easy to generate, everything in solution phase = higher efficiency
- Oligo pools can easily generate these Biotin-RNA or DNA bait molecules.



DNA data storage

According to leading manufacturer's of digital data mediums, the world's production capacity of traditional memory devices cannot keep pace with the increase in storage demand



Using DNA as a data storage medium is an example of looking to nature for technical solutions. DNA serves this function in the natural world and has some strong advantages. DNA is relatively stable compared to current digital data storage devices and is far more dense in its information capacity.

Size of world's 2013 data output.

Massive Attack's Mezzanine Reissued as **DNA** Spray Paint
SPIN - Oct 21, 2018
The press release also includes a statement from Dr. Robert Grass of Zurich's **TurboBeads**, who describes the process of encoding the **DNA**:.

Massive Attack releasing **DNA**-encoded Mezzanine in spray paint can
Consequence of Sound (blog) - Oct 19, 2018
Massive Attack releasing **DNA**-encoded Mezzanine in spray paint can ... A Dr. Robert Grass of Zurich's **TurboBeads** company explained the ...

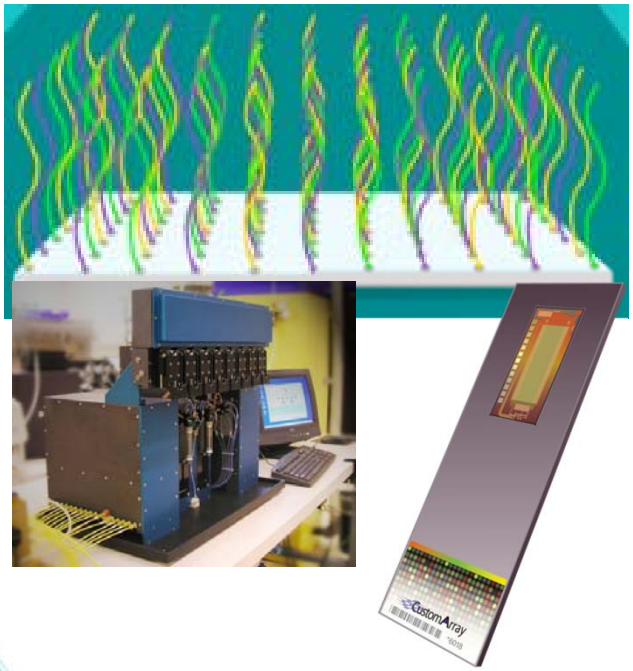
Massive Attack's Mezzanine "remastered" in aerosol spray can format
Highly Cited - FACT - Oct 19, 2018

View all

While these are two fun examples that CustomArray has contributed oligos for, there may be a strong case for using DNA as an archival storage medium

GenScript's oligo pool service

- ✓ **Maximum screening efficiency with >99% sequence coverage rate**
  - Our integrated platform can deliver every sequence in your order.
- ✓ **Low batch variations between oligo pools**
  - More confidence in your results when using multiple oligo pool batches.
- ✓ **Flexibility for your application**
  - Two chip sizes to create any pool size to meet your experimental needs.
- ✓ **No sequence restrictions or minimum order required**
- ✓ **Industry-leading turnaround time,** delivery as fast as 5 business days



Thank you!

For questions, please visit:
https://www.genscript.com/precise-synthetic-oligo-pools.html

or email:
kimberlya@genscript.com
oligo@genscript.com