SAND2011-2331C

Automated nucleic acid library preparation for sequence-based unknown pathogen detection

Hanyoup Kim, PhD

Sandia National Laboratories

Livermore California

LDRD
LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Overview

- **The emergence of unknown pathogens**
- **Approach to detecting unknown pathogens from clinical samples using next generation DNA sequencing**
- **Automated Molecular Biology (AMB) Platform**
  - Preparing Nucleic Acid libraries (from complex mixtures)
    1. Digital Microfluidic Hub architecture
    2. Transposase-mediated NA prep and barcoding
- **Conclusions**

Sandia National Laboratories

Novel pathogens increasingly threaten national security & public health

- **Factors promoting pathogen emergence:**
  - Increase encroachment on wildlife habitat
  - Increased population density, international travel & trade

- **Factors enabling pathogen engineering:**
  - Greater knowledge of pathogenicity & human biology
  - More, better, & cheaper tools for modification, synthesis, & evaluation of biological agents, including pathogens
  - Global dispersion of biological materials, knowledge, technology, & expertise

San Francisco Chronicle, Dec 20 2009

Our mission is to provide solutions to the most challenging biological analysis problems that threaten our national security.

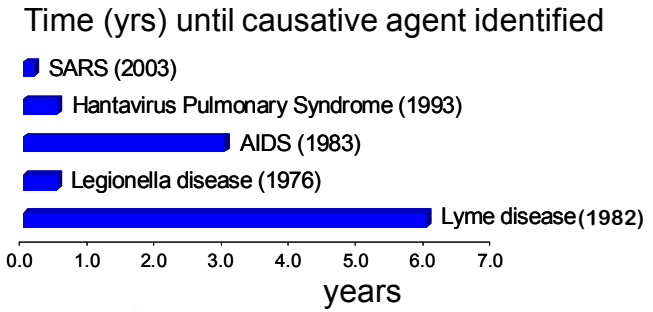

New York Times Feb 10, 2010

Sandia National Laboratories

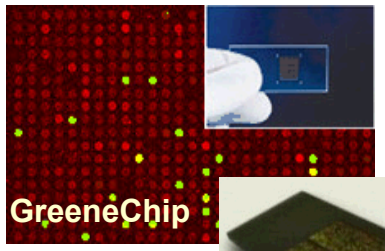
We need new tools for rapid identification & characterization of novel pathogens

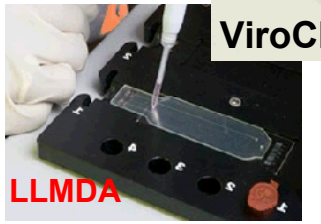
- Outbreak dynamics are often measured in ***days to weeks***.

- Identification of a novel causative agent by conventional methods can take ***months to years***.



- Modern probe-based methods are fast, but are often confounded by novel pathogens.
  - Pathogens can escape detection
  - Unanticipated features giving false negatives
  - Unusual profiles can be difficult to interpret





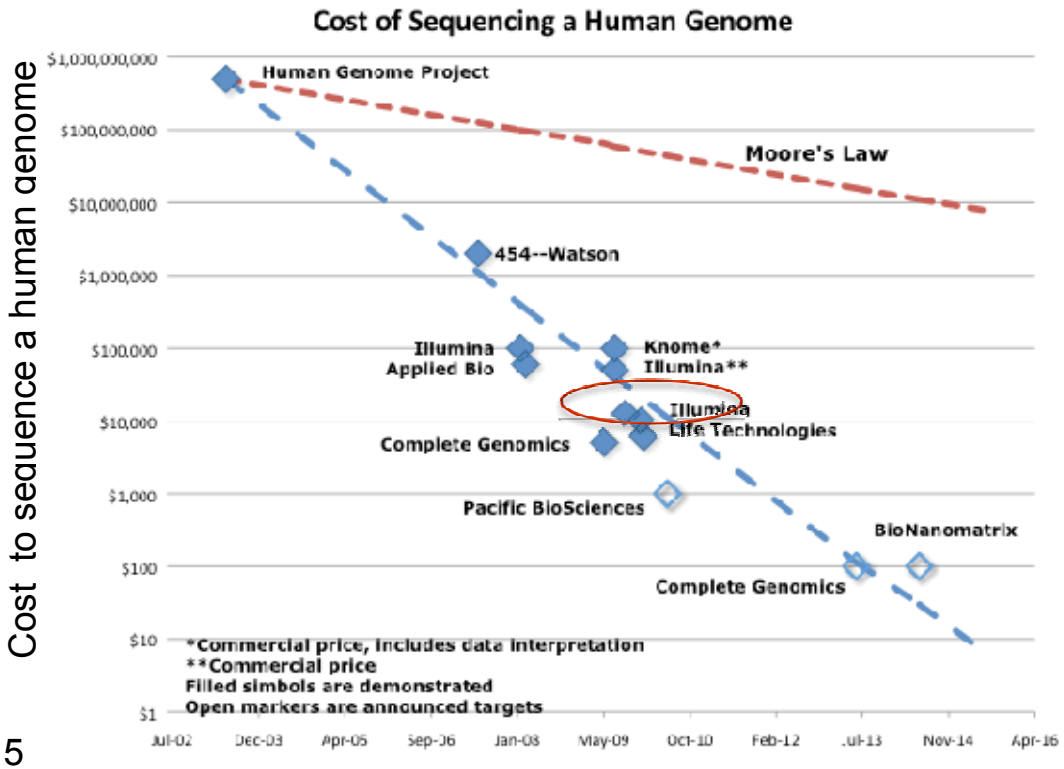
Sandia National Laboratories

Next Generation Sequencing is a transformational technology for pathogen characterization

- 150-200 Gb (2 x 100 bp run)
- Two human genomes (30 x coverage) in a single run under $10,000 per sample

http://www.illumina.com/systems/hiseq_2000.ilmn
ftp://ftp.ncbi.nih.gov/genbank/gbrel.txt

http://www.illumina.com/systems/hiseq_2000.ilmn



http://www.flickr.com/photos/doe_jgi/3876606040

…but DNA sample prep is primarily a benchtop process

Sandia National Laboratories

Brute-force NGS of clinical samples can enable discovery of novel pathogens

Disease	Sample	Novel Agent Detected	Total Reads	Hits on Agent	Reference
Merkel cell carcinoma	tumors	"Merkel cell polyomavirus"	395,734	2 (**0.00005%**)	cience 319:1096 '08
organ transplant related fatality	serum & organs	"Dandenong" arenavirus	103,632	14 (**0.014%**)	N Engl J Med 358:991 '08
pediatric gastroenteritis	feces	"human klassevirus "	937,935	849 (**0.09%**)	Virol J 6:82 '09
pediatric influenza-like illness	nasopharyngeal swabs	"human enterovirus type 109"	20,825,810	119 (**0.0006%**)	Virol 84:9047 '10

Deplete non-informative NA to improve efficiency of NGS analysis

Sandia National Laboratories

The Challenge: Develop a new approach to rapidly characterize unknown bioagents

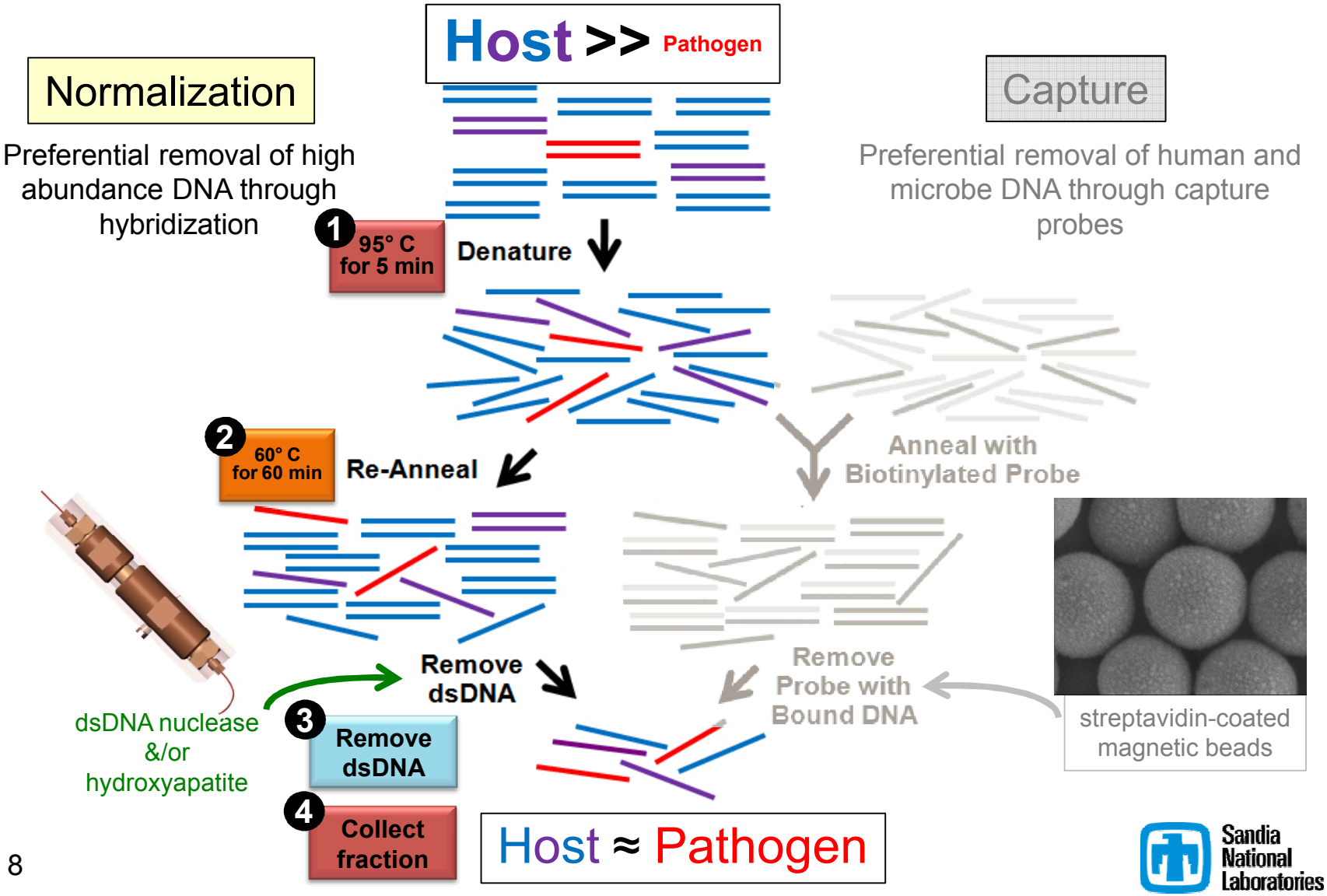
Rapid Threat Organism Recognition (RapTOR) system

- **Goal:** Efficient analysis of pathogen nucleic acids (NA) in clinical samples *via* targeted Next Generation Sequencing (NGS)
- **Key advance:** Automated microfluidic platform to enable molecular suppression and NA preparation to improve signal-to-background (pathogen-to-host) NA ratio in samples

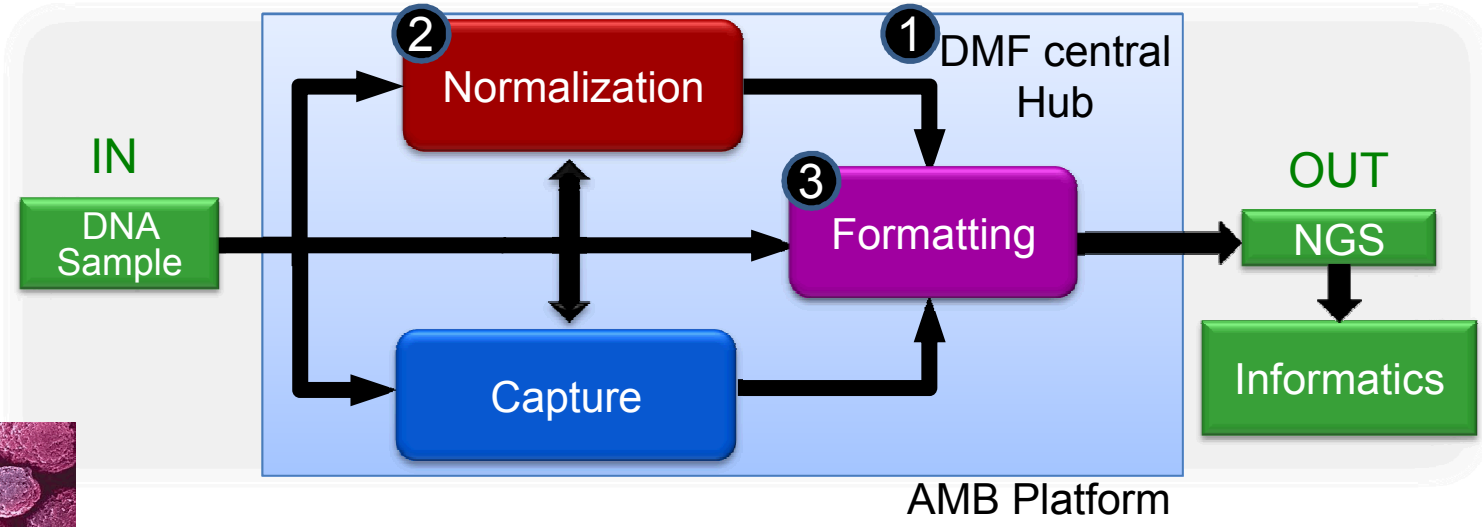
Drivers for our approach

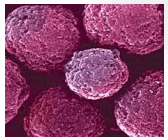
- Identify and characterize unknown pathogens in a timeframe compatible with rapid disease outbreak detection and response
- No prior knowledge of a pathogen or culturing of organism; (deep genomic sequencing)
- Automate the nucleic acid processing for operation at federal and state-wide laboratories

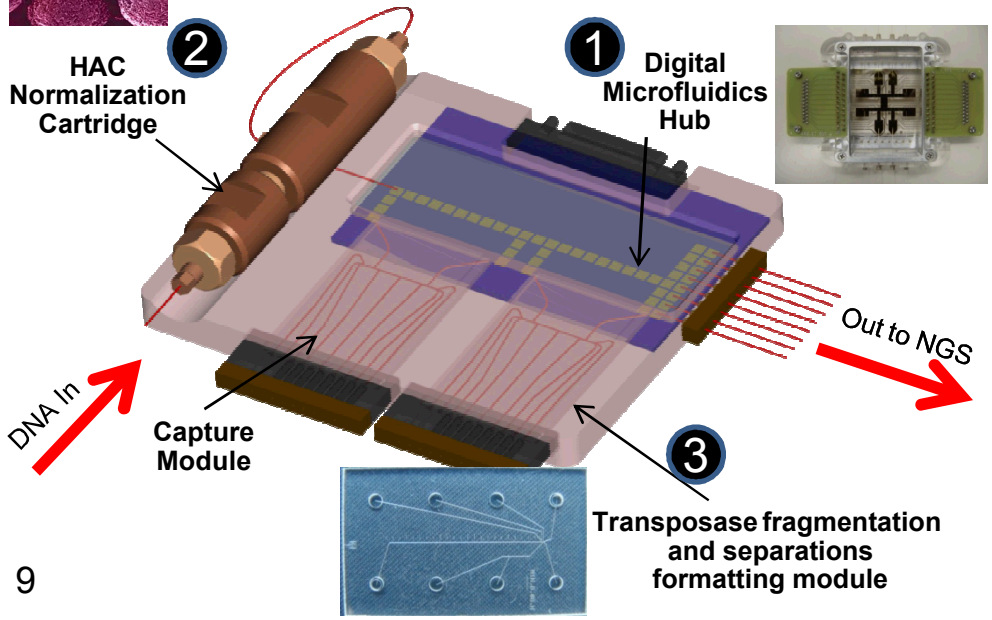
We are focusing on complementary suppression methods for depletion of host NA



AMB platform integrates suppression & library prep methods into a flexible architecture





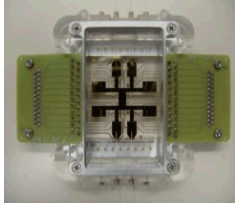


Gen 1 requirements:

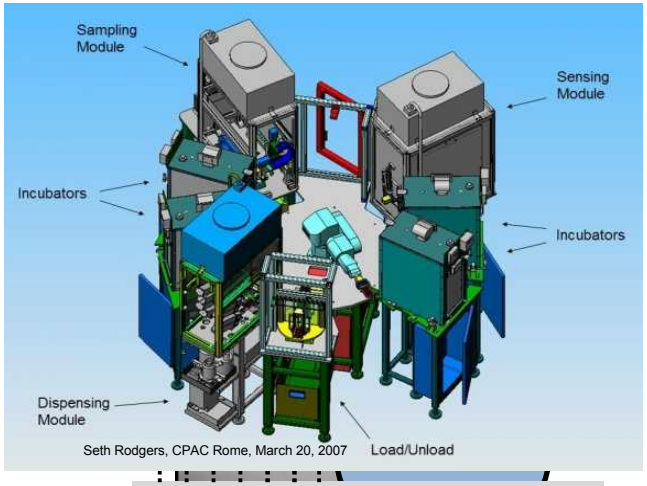
- Normalization + Formatting
- Handle ng quantities of gDNA and cDNA
- 10-100 fold suppression
- Semi-automated

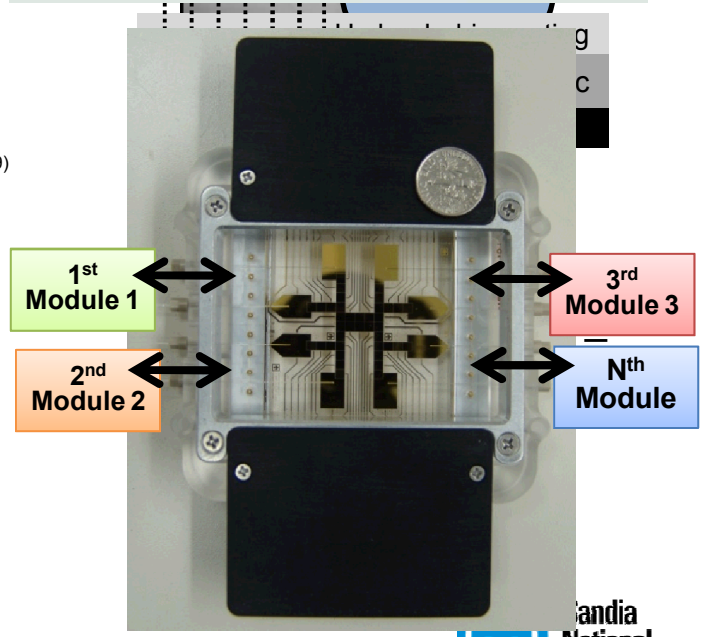
Sandia National Laboratories

Core architecture of the AMB platform for NA processing is the Digital Microfluidic (DMF) Hub



- Use droplets as sample cargo containers
  - Operated "digital" fashion (virtual tubes or microreactors)
  - Nanoliter to microliter in volume
  - Merge, mix, split (virtual pipetting)
- Based on principles of electrowetting-on-dielectric (EWOD) and dielectrophoresis
  - voltage is applied to electrode pads in an addressable 2-D array on glass substrates.
- Pollack and Fair at Duke University
  - Oil/water system –PCR in DMF (Hua et al. *Anal Chem* 2010)
- Aaron Wheeler's group at U. of Toronto
  - Air/water system—cell-based microculturing (Bluovak et al, Lab Chip 2009)
- DMF as a central hub for interfacing multiple lab-on-a-chip sample processing modules through droplets
  - Advantage
    - Flexibility and spatial manipulations of droplets
    - Modularity and temporal resolution of continuous-flow microchannel devices
    - Overcome world-to-chip interface difficulties
      - Sample volume mismatch & timing



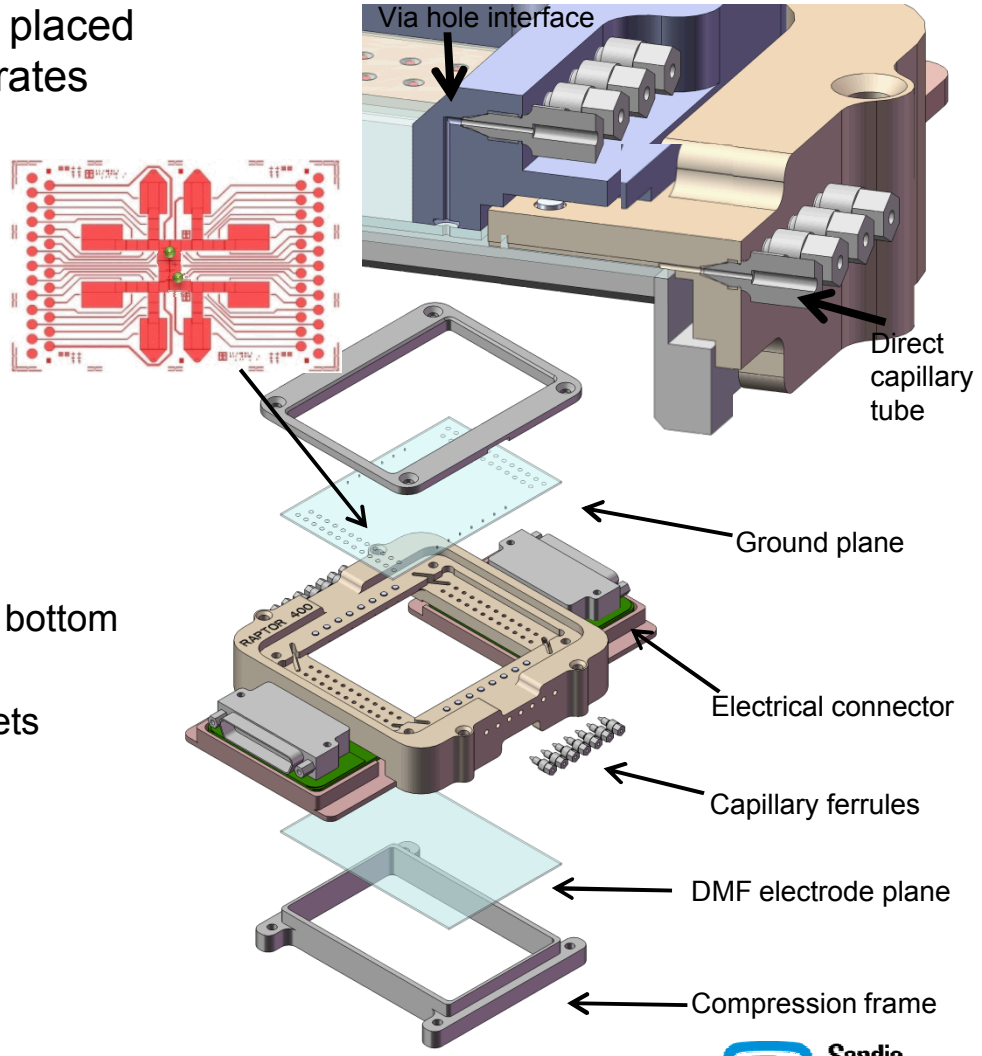


Sandia National Laboratories

DMF hub platform uses capillary tubes for reliable droplet to module interface

- Teflon-coated, 150 um o.d capillaries placed in gap between the DMF glass substrates
- Top-side through via holes

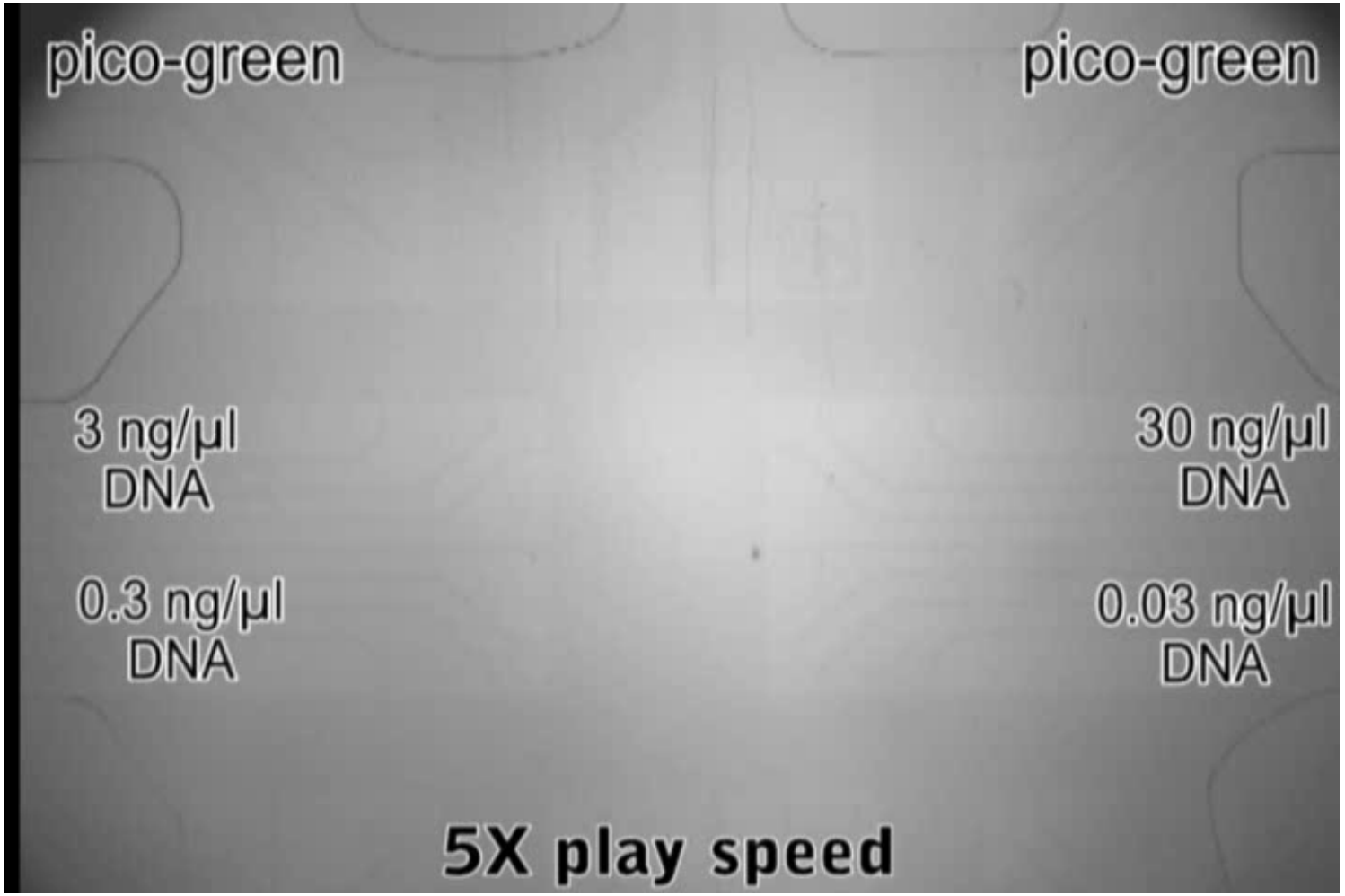
- **Custom-fabricated manifold**
  - plug-and-play installation of the DMF substrates
  - electrical contact for 46 actuation pads
  - Up to 14 capillary interface ports
- **DMF substrates**
  - ITO-coated glass for complete top and bottom optical access
  - 185-um gap size microliter-sized droplets using ~2x2 mm pads
- **Programmable interface**
  - Manual keyboard keys
  - Script routines



Bartsch, Kim and Patel, *Sensor and Actuators A*, 2011 in prep.

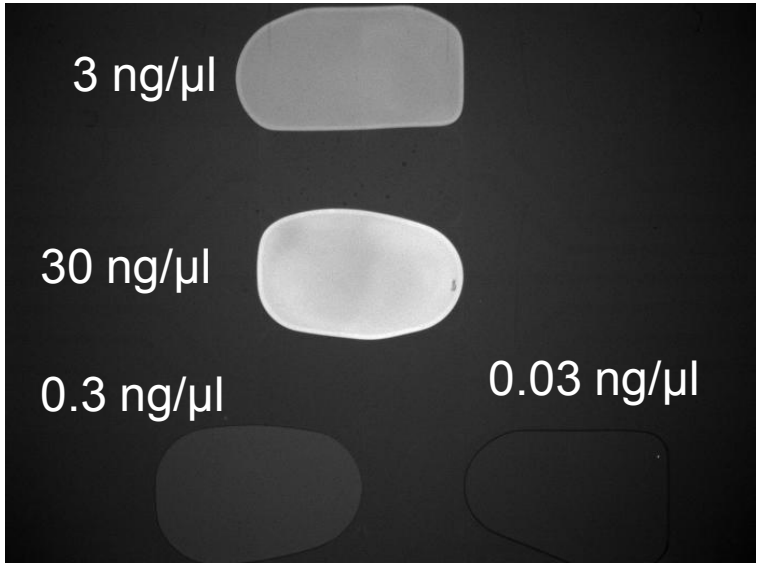
Sandia National Laboratories

Single assay using picogreen: 10^3 fold DNA concentration quantification

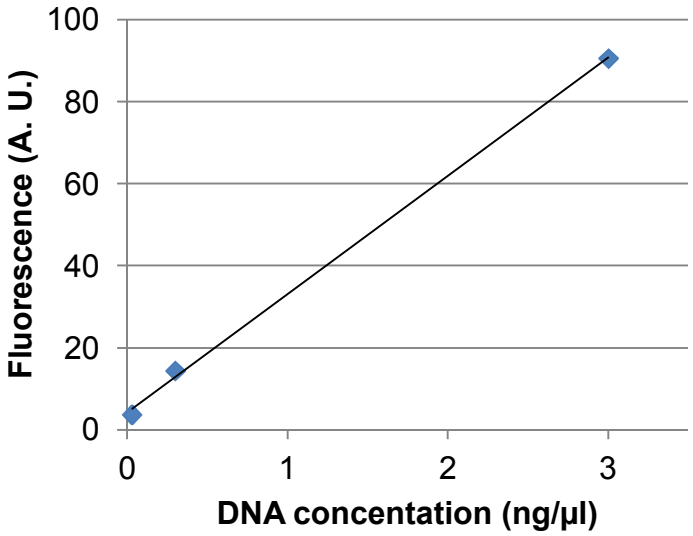


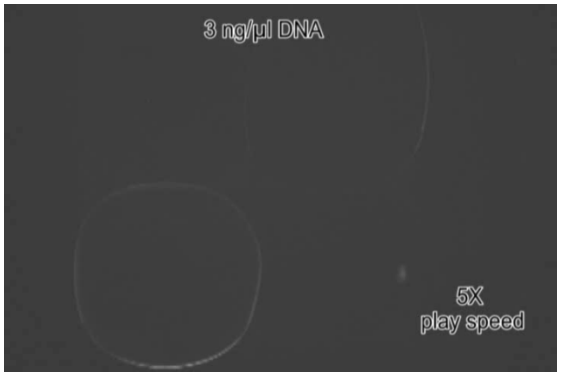
Sandia National Laboratories

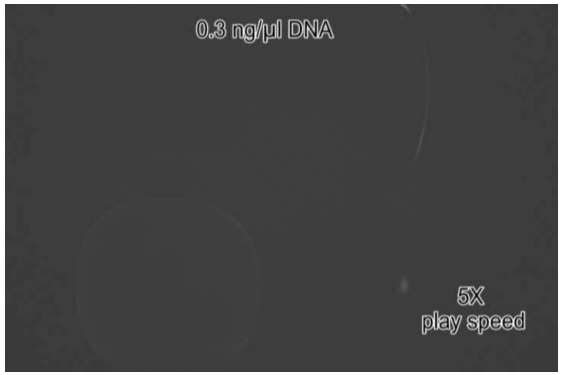
Single assay using picogreen: 10^3 fold DNA concentration quantification



Fluorescence image





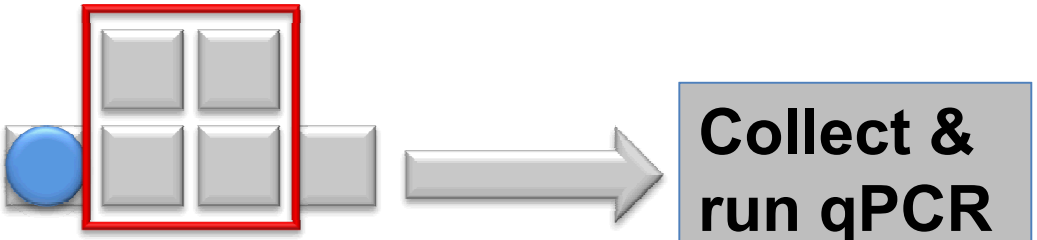


Sandia National Laboratories

Carry over well controlled (10^6-fold dilution)

Sequence of Droplet introduced into 10-pads carryover test zone

1. **Control buffer drop for baseline**
2. **cDNA droplet for contamination**
3. **Control buffer droplets to assess carry over**
4. **Recover droplets and run qPCR (beta actin)**



Carry over test zone

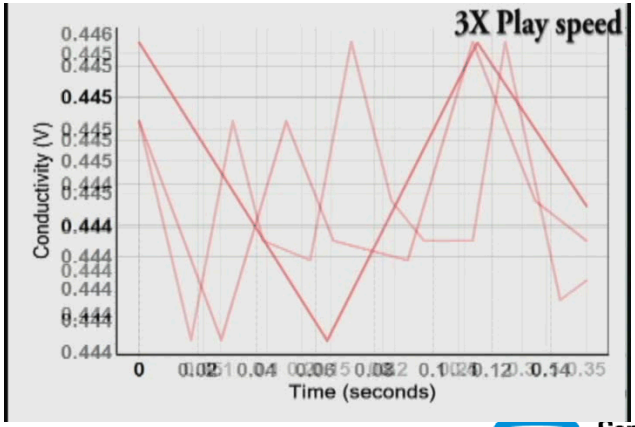
	Buffer drop (Baseline)	cDNA Drop	1st Buffer Drop	2nd Buffer Drop	3rd Buffer Drop
Ratio to cDNA	$<10^{-8}$		$\sim10^{-6}$	$\sim10^{-7}$	$\sim10^{-8}$

Only 1 out of ~1,000,000 DNA molecules left behind and picked up by subsequent buffer droplet on DMF

Sandia National Laboratories

Example of a capillary to DMF interface for discretizing eluting peaks

- Teflon-coated, 150 um o.d. capillaries placed in gap between the DMF glass substrates
- Dark-field with epifluorescence made possible by transparent DMF substrates
- Contactless conductivity sensor sheathed around the capillary to detect conductivity changes to enable switching and fraction collection
- 4 µL Fluorescein in 130 mM Na_3PO_4

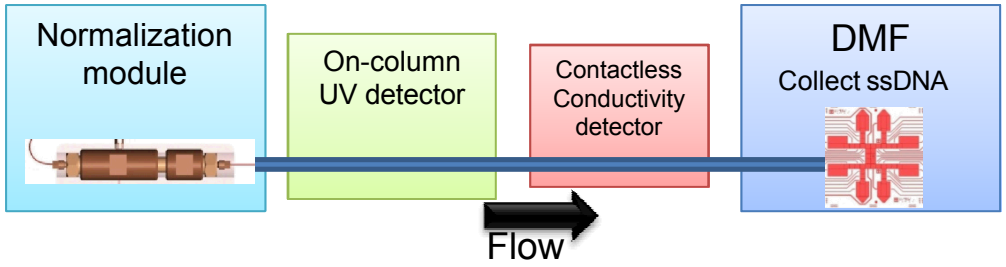


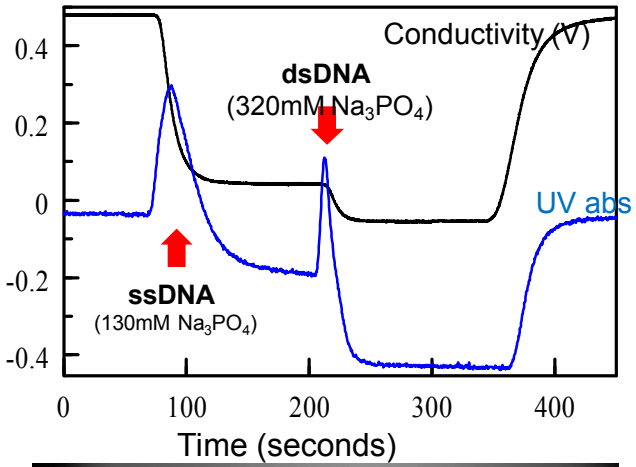
Kim & Patel, 2010

Sandia National Laboratories

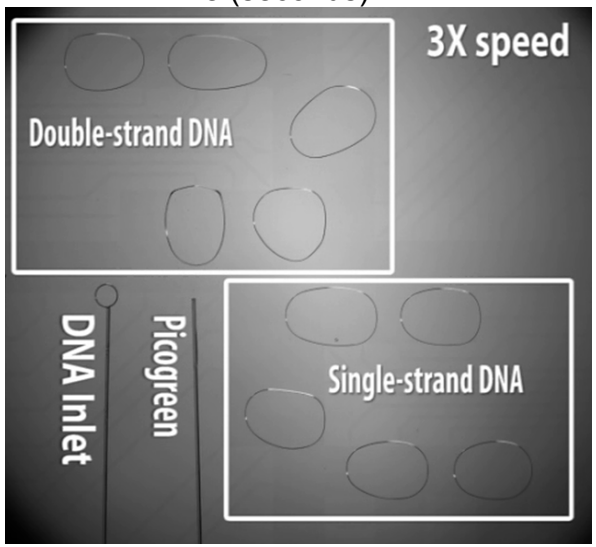
Demonstration of HAC normalization module integration w/ DMF using a capillary interface

- **Normalization**
  - Removal of high abundance DNA through selective hybridization
  - Denature, anneal, separate, collect ssDNA





- ssDNA labeled with cy3 (100 bp)= **low abundance**
  - 160 mM Na_3PO_4 buffer
- dsDNA: DNA ladder >100 bp = **high abundance**
  - 330 mM Na_3PO_4 buffer

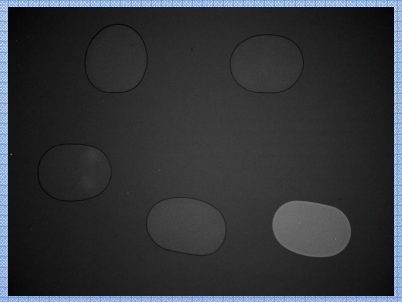
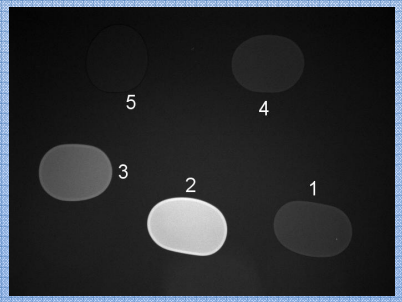
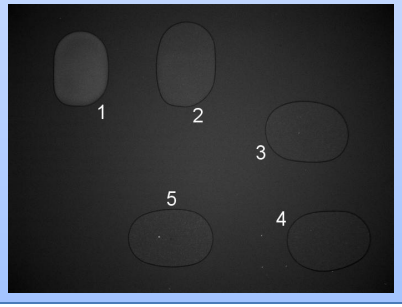


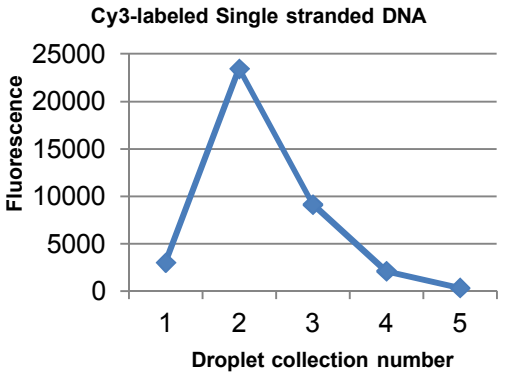
- Second downstream capillary dispenses 100 nL of picogreen for on-demand labeling on the DMF device
- Conditions:
  - 2 µL HAC column cartridge
  - Flow rate ~15 µL/min
  - Step gradient 10mM: 160mM: 320mM Na_3PO_4

Kim, Vandernoot and Patel *Anal Chem, 2011 in prep*

Sandia National Laboratories

Demonstration of HAC normalization module integration w/ DMF using a capillary interface

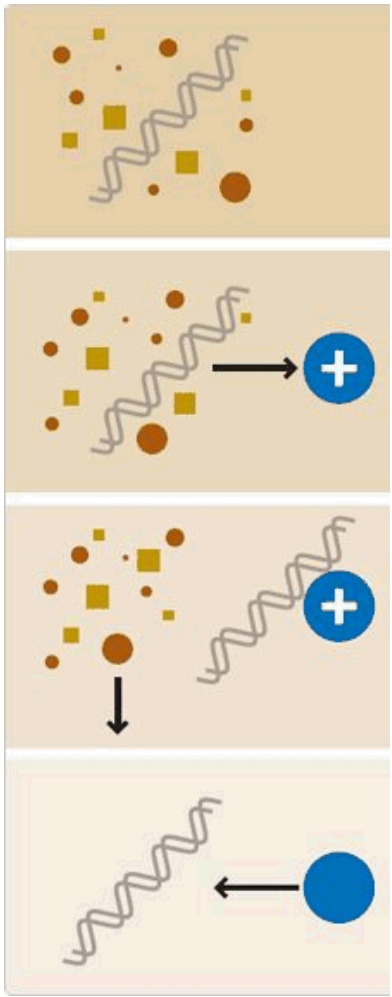
	Picogreen	Cy3
1st eluting peak		
2nd eluting peak		



ssDNA eluting peak is discritized into droplets on the DMF platform

Sandia National Laboratories

Magnetic beads assay performed on DMF captures DNA effectively



DNA + Contaminants

DNA with contaminants
Ex) Excessive salts, enzymes

DNA binding

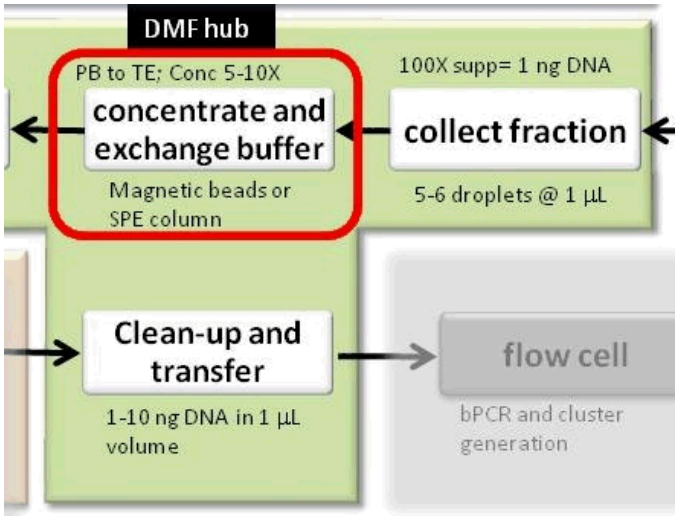
Purification buffer lowers pH
Beads added take positive charges & bind DNA

Contaminants Removal

Collect beads pellet
Wash & collect beads pellet

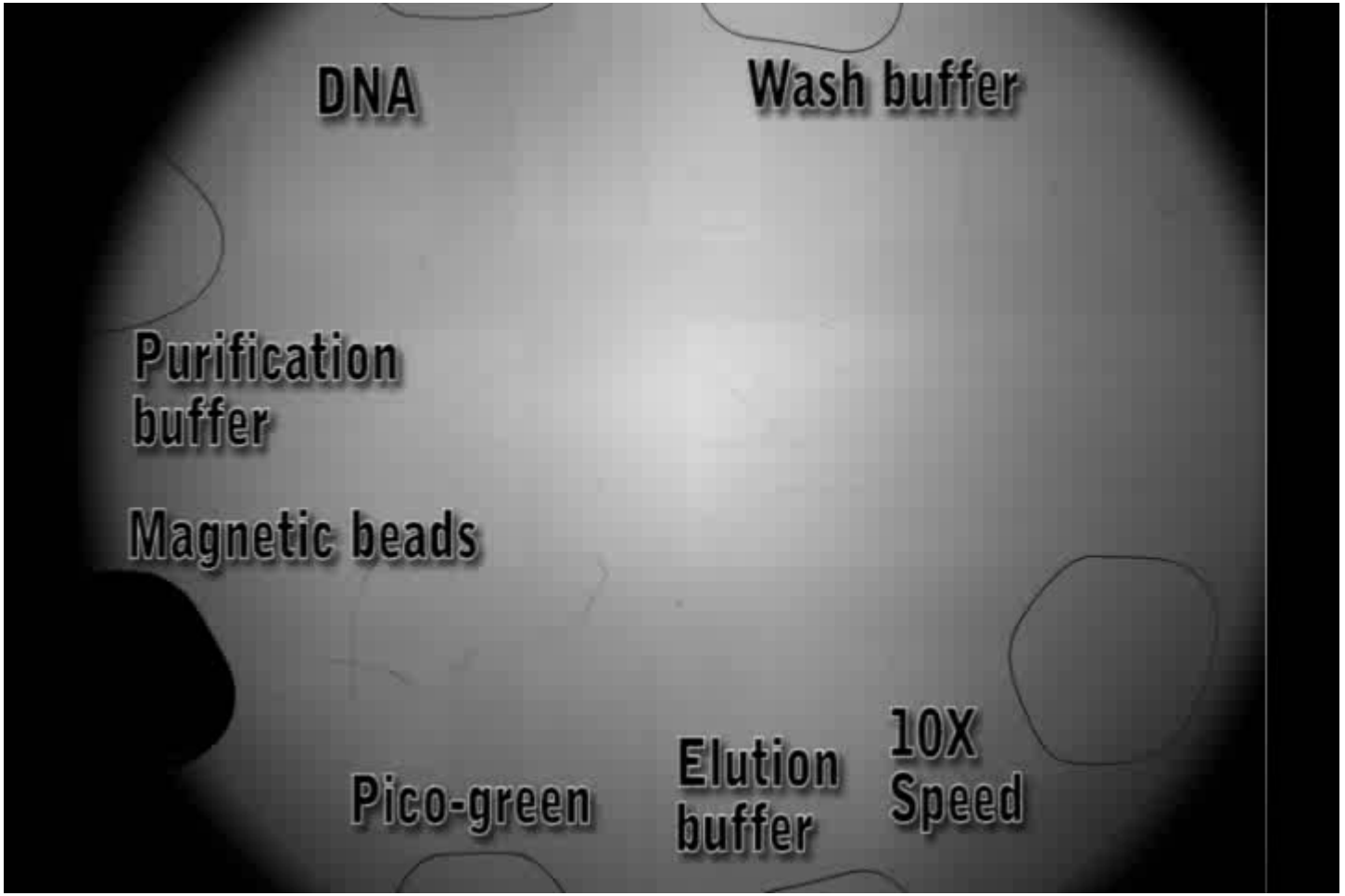
DNA Elution

Elution buffer raises pH
Collect elution buffer only



Sandia National Laboratories

Magnetic beads assay performed on DMF captures DNA effectively

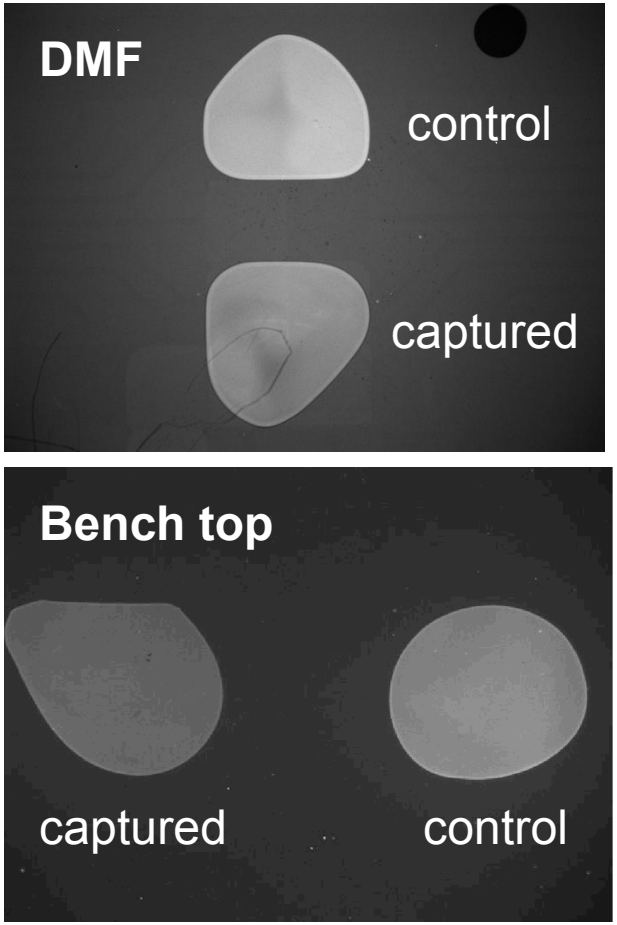


Kim, 2010

Sandia National Laboratories

Magnetic beads assay performed on DMF captures DNA effectively

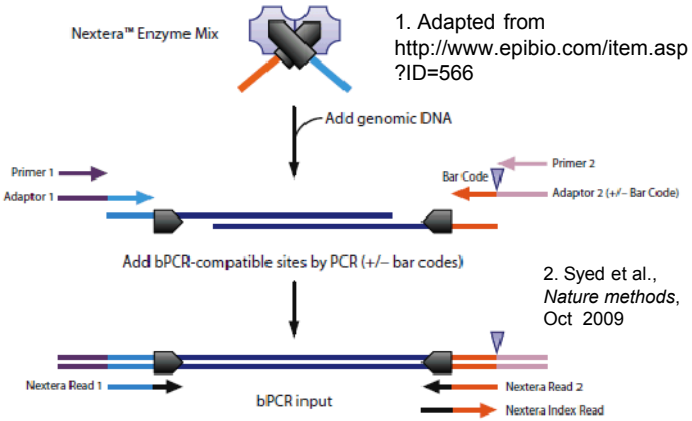
- Assay on DMF hub shows comparable capturing efficiency of DNA to Bench top process
- Reproducibility and efficiency studies show 80 ± 4.8% in 3 independent runs (AMPpure® xp beads assay)
- **Near future goals:**
  - Improve efficiency with capillary interface mixing
  - Use biotinylated magnetic beads for exome capture probes for orthogonal suppression technique



Sandia National Laboratories

Transposase-mediated fragmentation and ligation simplifies library preparation

- **Single tube reaction,**
  - **Fragments and ligates Illumina adapters DNA (w/ barcodes)**
  - **Drastically reduces number of steps and preparation time**



1. Adapted from http://www.epibio.com/item.asp?ID=566

2. Syed et al., *Nature methods*, Oct 2009

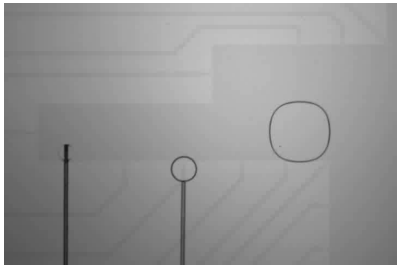
Illumina Protocol (µg)	min	Nextera Transposase-mediated(~50 ng)	min	AMB-adapted Transposase (~pg)	min
Fragmentation	30	Add Nextera™ Enzyme Mix to DNA	5	React transposase + DNA on DMF device with thermal capillary reactors	5
Collection	15				
Concentration	15				
Size Selection	60				
End-Repair	60			**Quantitation of DNA before and after PCR using DMF interfaced chip electrophoresis**	**15**
Clean-Up	15				
A-Tailing	30				
Adaptor Ligation	60				
Clean-Up	15	Clean-Up and size selection	60	Bead-based clean-up and size separation	~20
Benchtop PCR (Enrichment)	~60	Benchtop PCR	60	DMF interfaced PCR	~20 min

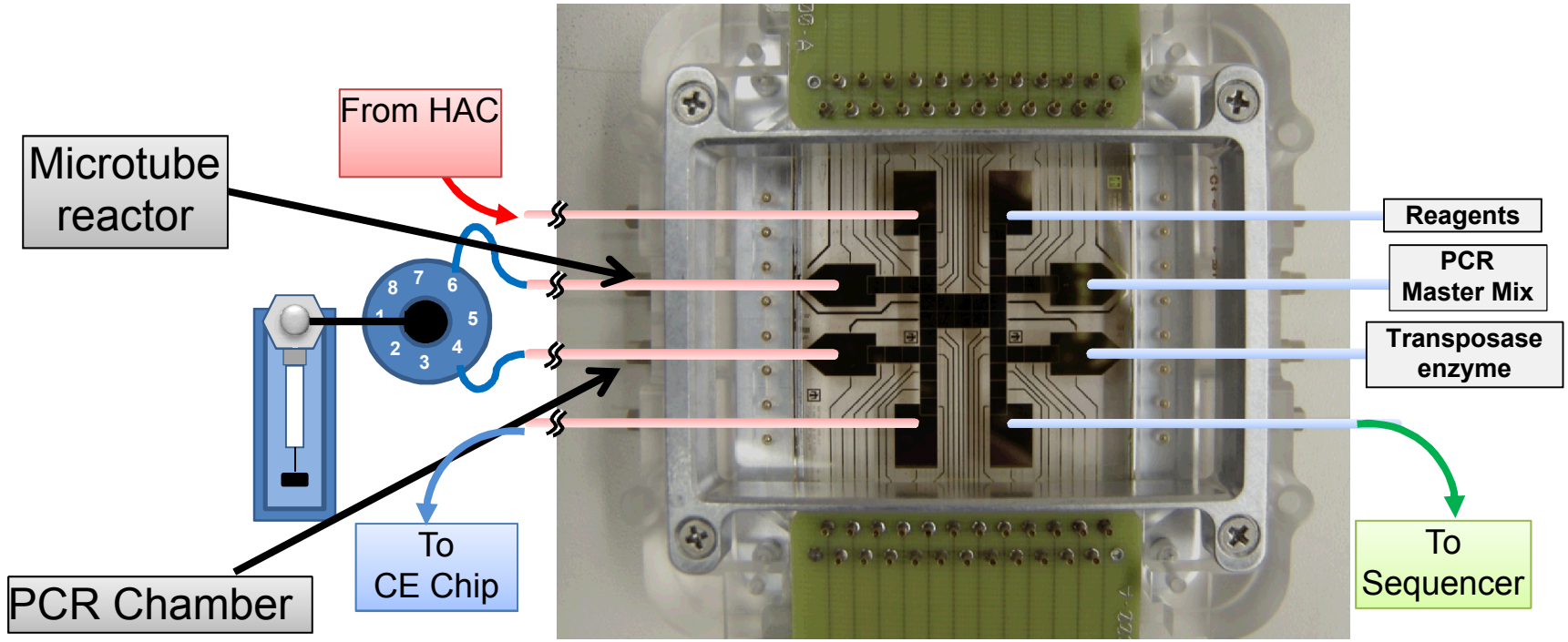
~ 6 hrs **~ 2 hrs** **~1 hour**

National Laboratories

Transposase-mediated reactions at pg-scale can be adapted to DMF device through interfaced modules

1. Mix enzyme with DNA
2. React at 55° C for 5 min
3. PCR amplify
4. Clean-up and size DNA
5. Quantitate DNA through separations

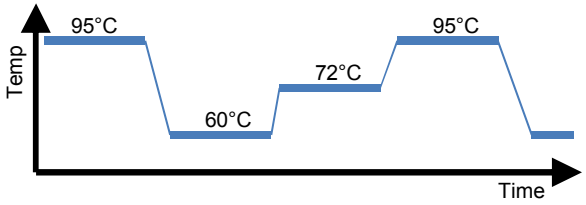


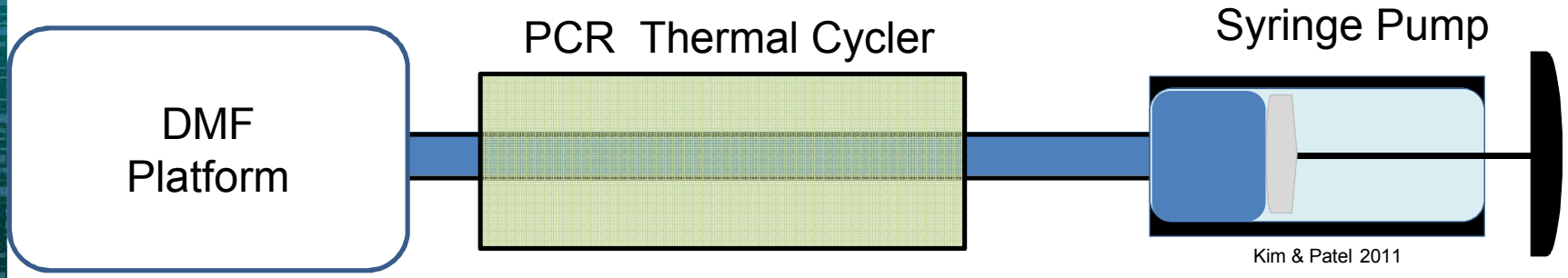


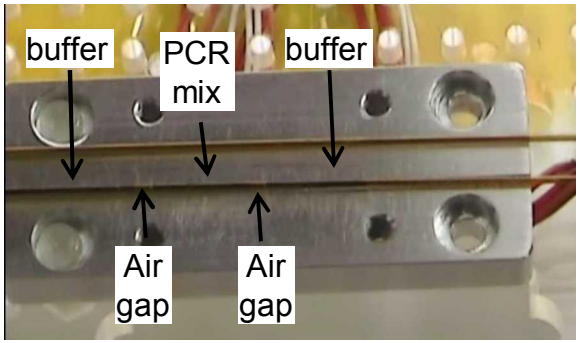
Sandia National Laboratories

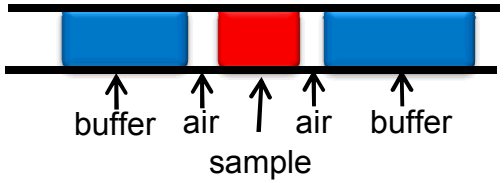
Off-chip PCR module can be interfaced with the DMF platform

- Sample bolus with air separators is temperature cycled for amplification
  - Scalable from nL to µL (change tubing diameter)
  - Interface with DMF and droplet reagents
  - Disposable (inserted tube)
  - Multiplex using multiple capillaries





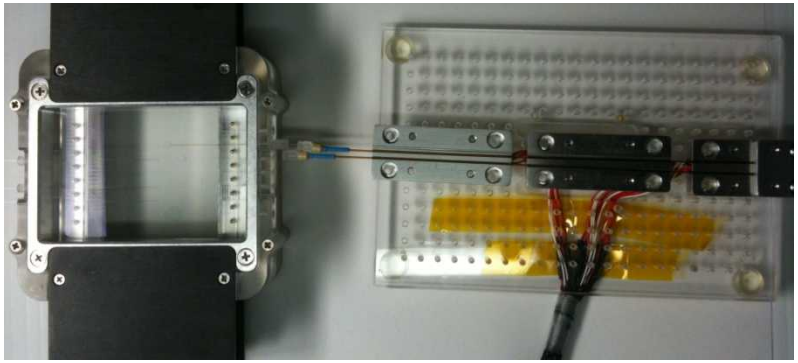




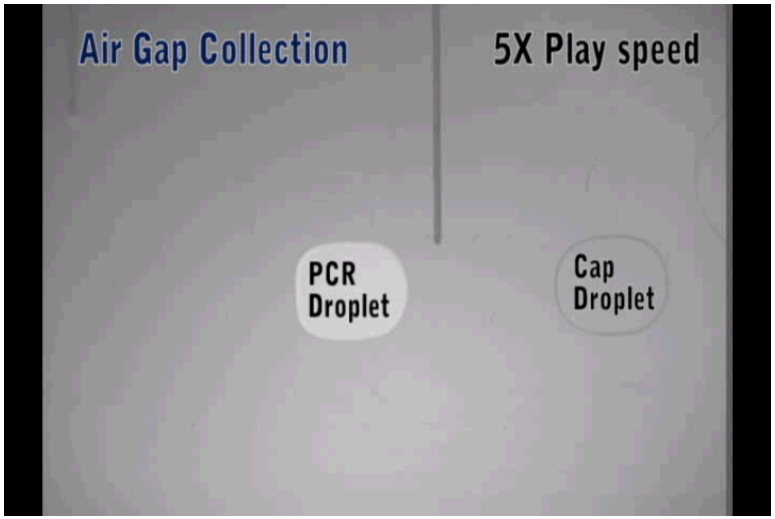
Sandia National Laboratories

Fluidic Interface with DMF for PCR has been demonstrated

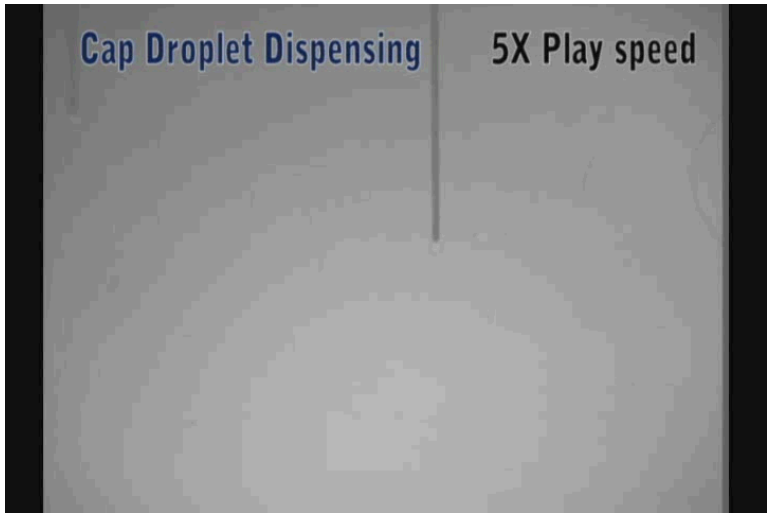
- 2-6 µL PCR reactor
- Repurposed for enzyme-based thermal reactors
  - Transposase @ 55 deg for 5 min
- Demonstrating proof-of-principle



PCR interfaced to DMF



Load PCR reactor

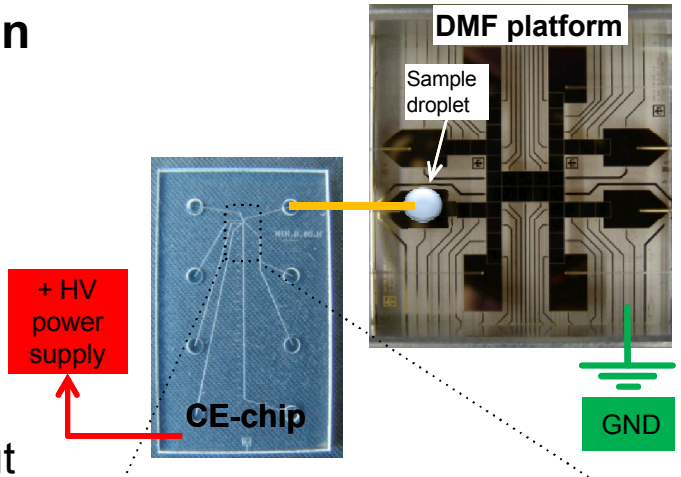


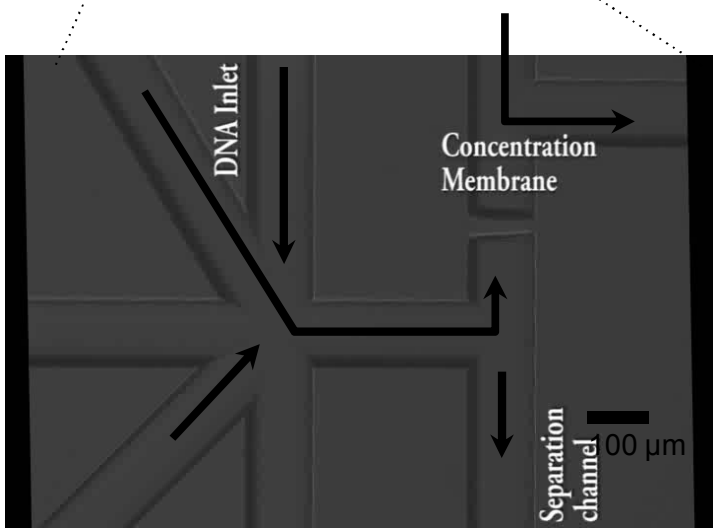
Collect PCR products Patel ,2010

Sandia National Laboratories

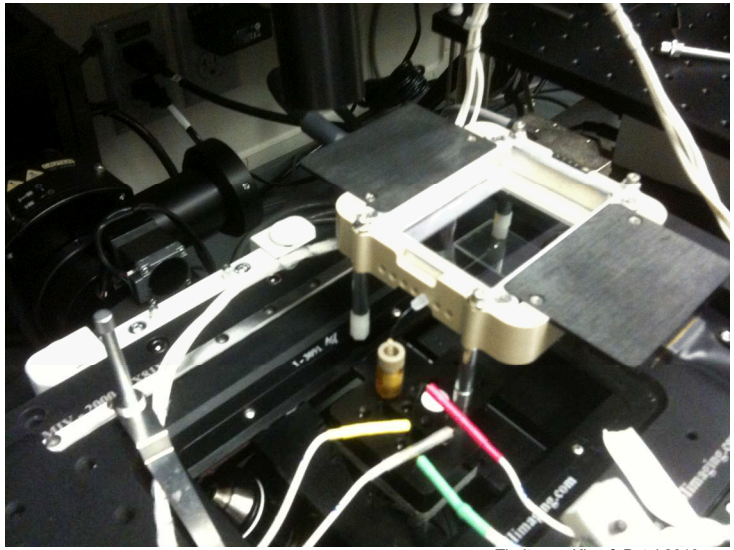
Quantitative analysis of fragmentation and PCR products from the DMF droplet with CE

- **Electrophoresis of DNA from a droplet on DMF device to a nanoporous membrane**
  - Use DMF ground plane to complete circuit
  - Sample small fraction of droplet for separation analysis
- **Nanoporous membrane**
  - Photo-polymerized *in situ* with a UV laser
  - Traps DNA to concentrate at the surface but allow ions to pass





Thaitrong and Patel, 2011



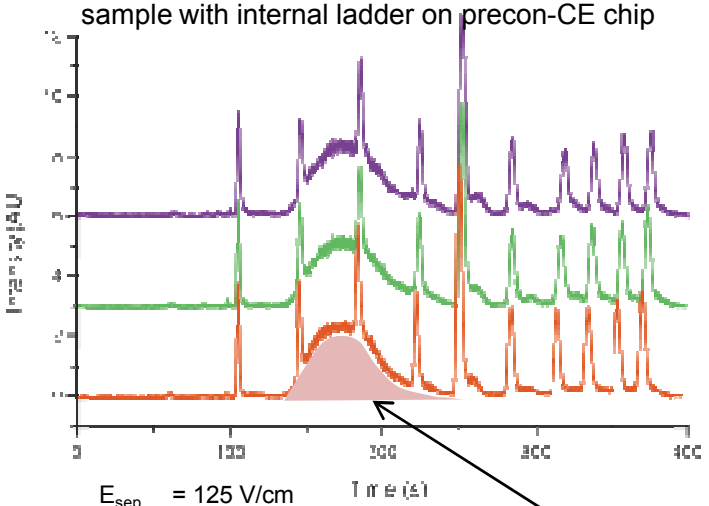
Thaitrong, Kim, & Patel 2010

Sandia National Laboratories

Preconcentration integrated with CE-chip enables sensitive LIF detection of NA reaction products

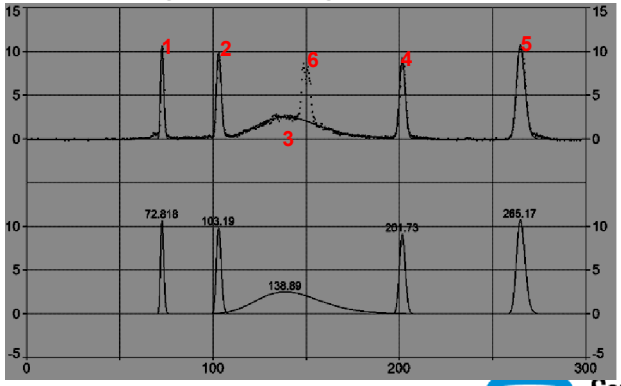
- Preconcentration minimizes effect of diffusion and dilution of DNA sample
- DNA stacks at the membrane interface a well-defined injection plug at the separation column
- Preconcentration improves our detection limits (10-100-fold)
- Add an internal ladder to quantitate our reaction products
  - Calculated using peak area of co-injected ladder as an internal standard

Separation of transposase-mediated fragmentation sample with internal ladder on precon-CE chip



E_{sep} = 125 V/cm
T_{conc} = 100 s
SYBR green intercalating dye

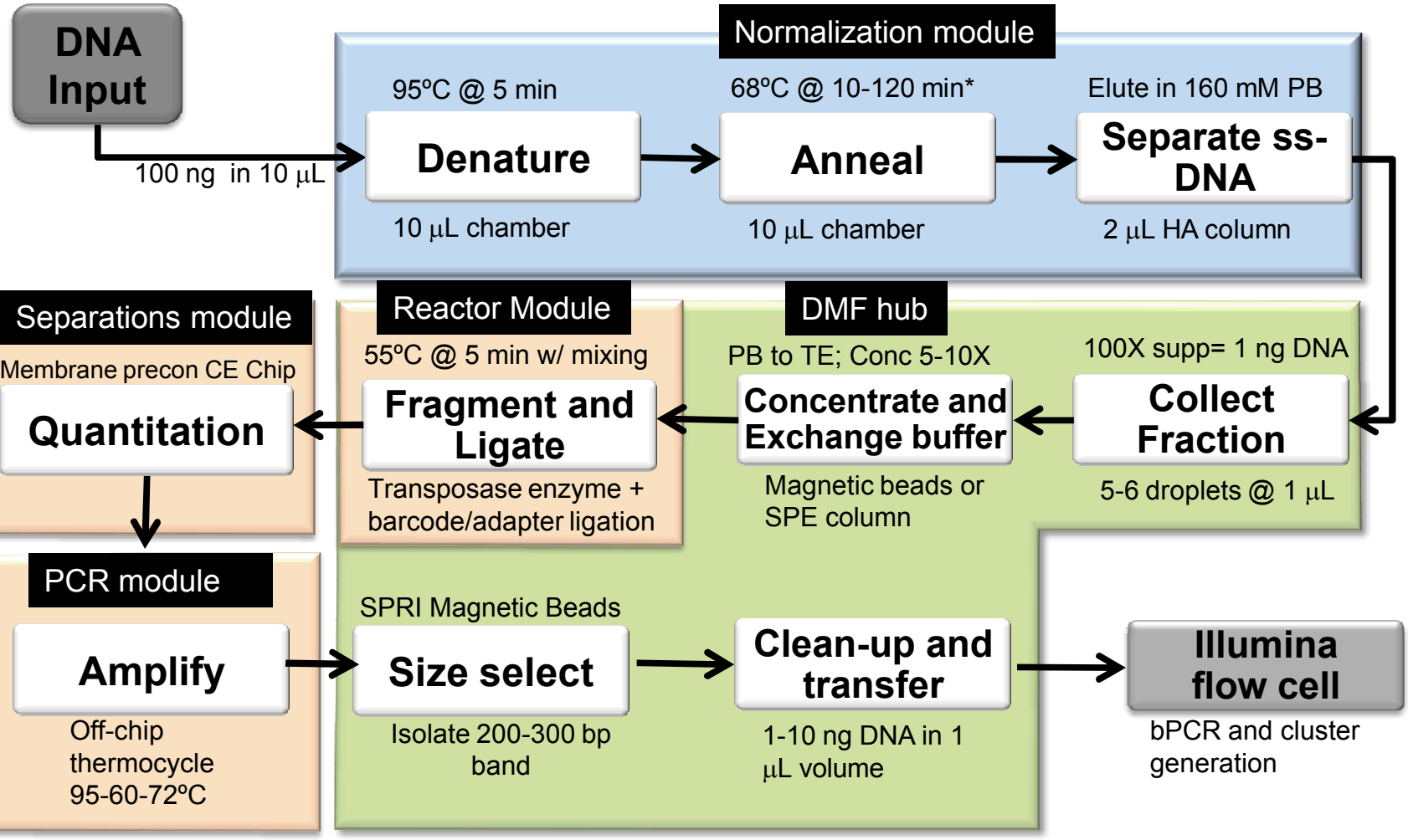
1 ng/µL ± 165 pg/µL (n = 3)



Thaitrong, 2010

Sandia National Laboratories

NA prototype process flow diagram for normalizing + NA sequencing preparation



Sandia National Laboratories

Acknowledgments

- **Automated Molecular Biology Team**
  - Team lead: Kamlesh Patel
  - **Hanyoup Kim, Numrin Thaitrong**, Robert Meagher, **Victoria VanderNoot**, Conrad James, Carlton Brooks,
  - Engineering Team **Michael Bartsch**, **Ron Renzi**, Jim He, Jim Van De Vreugde, Ron Renzi, Mark Claudnic
- **Host Pathogen Molecular Biology Team**
  - Team Lead: Steve Branda
  - **Stan Langevin, Zach Bent**, Sadie LaBauve, Bryan Carson, Julie Kaiser, Pam Lane, Bryce Ricken, Deanna Curtis
- **Data knowledge and Discovery Team**
  - Team Lead: Joe Schoeniger
  - Milind Misra, Kelly Williams, Amy Powell, Chi-Chi May
- **Project Management and PI:**
  - Duane Lindner, Malin Young, and Todd Lane

Questions

Thank You

Sandia National Laboratories