



SureSelect The Leader in Target Enrichment

Genomic Solutions Division

Rangasamy Elumalai Ph.D.
Applications Scientist
07-14-2016

1

SureSelect Technology

2

Whole Exome Sequencing

3

FFPE Samples

4

Custom DNA Panels

5

Methyl-Seq

6

Halo-HS

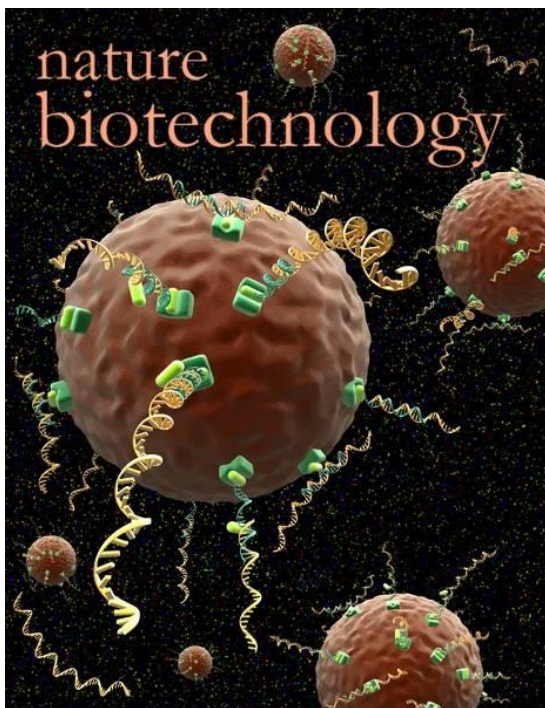
7

Custom Oligo Library Access



SureSelect Technology

Pioneer in NGS Target Enrichment



ARTICLES

nature
biotechnology

Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing

Andreas Gnirke¹, Alexandre Melnikov¹, Jared Maguire¹, Peter Rogov¹, Emily M LeProust², William Brockman^{1,5}, Timothy Fennell¹, Georgia Giannoukos¹, Sheila Fisher¹, Carsten Russ¹, Stacey Gabriel¹, David B Jaffe¹, Eric S Lander^{1,3,4} & Chad Nusbaum¹

Targeting genomic loci by massively parallel sequencing requires new methods to enrich templates to be sequenced. We developed a capture method that uses biotinylated RNA 'baits' to fish targets out of a 'pond' of DNA fragments. The RNA is transcribed from PCR-amplified oligodeoxynucleotides originally synthesized on a microarray, generating sufficient bait for multiple captures at concentrations high enough to drive the hybridization. We tested this method with 170-mer baits that target > 15,000 coding exons (2.5 Mb) and four regions (1.7 Mb total) using Illumina sequencing as read-out. About 90% of uniquely aligning bases fell on or near bait sequence; up to 50% lay on exons proper. The uniformity was such that ~60% of target bases in the exonic 'catch', and ~80% in the regional catch, had at least half the mean coverage. One lane of Illumina sequence was sufficient to call high-confidence genotypes for 89% of the targeted exon space.

The development and commercialization of a new generation of increasingly powerful sequencing methodologies and instruments¹⁻⁴ have lowered the cost per nucleotide of sequencing data by several orders of magnitude. Within a short time, several individual human

have been tested on target sets complex enough to match the scale of current next-generation sequencing instruments.

The first method, microarray capture^{9,12,13}, uses hybridization to arrays containing synthetic oligonucleotides that match the target

- Capture fragments with longest, most efficient 120-mer cRNA baits
- Probes can be designed to any regions of interest, samples can be multiplexed
- Easy to implement and compatible with validated automation solution

SureSelect Technology

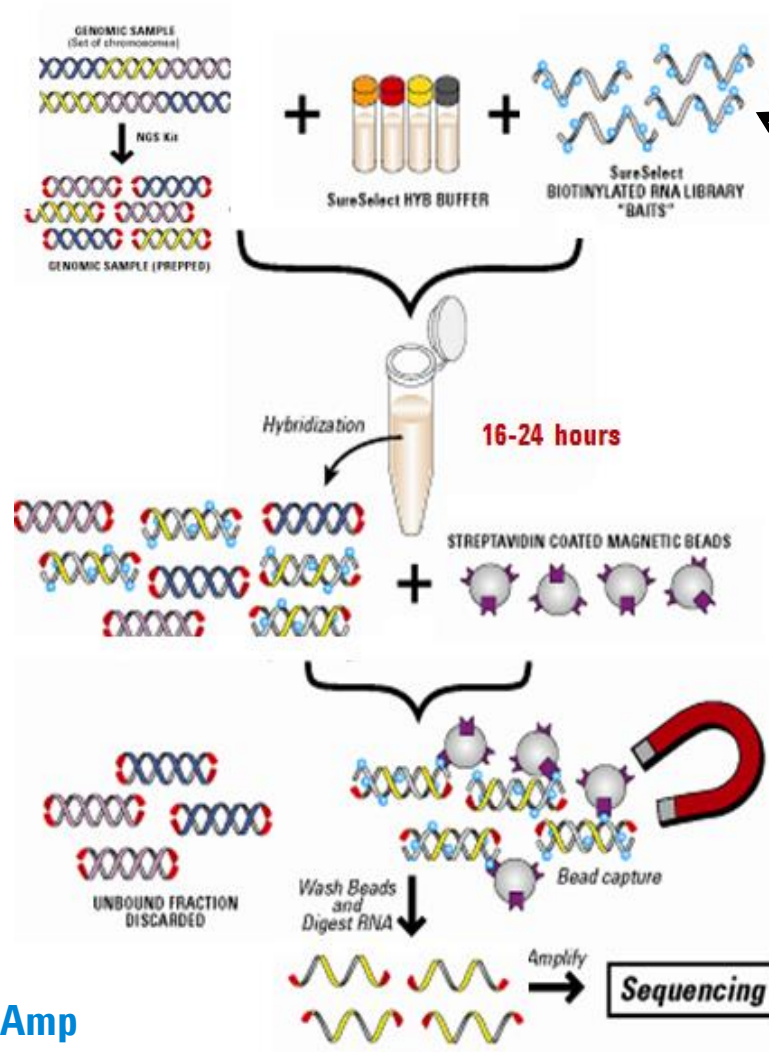
Simple Target Enrichment Workflow

Library prep
Hyb set-up

Hybridization

Capture

Wash / Elution / Amp



Baits:

- cRNA probes
- Long (120bp)
- Biotin labeled
- User-defined



>50 Mendelian Diseases uncovered by SureSelect

- Miller syndrome – USA - 2009
- TARP syndrome – USA - 2010
- Schinzel-Giedion syndrome - Netherlands - 2010
- Fowler Syndrome – Canada - 2010
- Terminal Osseous Dysplasia – Netherlands – 2010
- Hearing Loss – USA – 2010
- Perrault Syndrome – USA – 2010
- Kaposi sarcoma – USA – 2010
- Sensenbrenner Syndrome – Netherlands – 2010
- Hyperphosphatasia syndrome – Germany – 2010
- Kabuki syndrome - USA – 2010
- Van Den Ende-Gupta syndrome – Canada – 2010
- Neonatal Diabetes Mellitus – France – 2010
- Autoimmune lymphoproliferative syndrome – USA – 2010
- Familial Amyotrophic lateral sclerosis – USA – 2010
- Non-syndromic mental retardation – USA – 2010
- Osteogenesis Imperfecta – Germany/Netherlands – 2011
- Hajdu-Cheney syndrome – London /France– 2011
- Acne Inversa – China – 2011
- Leucoencephalopathy – Japan – 2011
- Taybi-Linder syndrome – France – 2011
- Ochoa syndrome – Saudi Arabia – 2011
- Spastic paraparesis - USA – 2011
- Distal Arthrogryposis – USA – 2011
- Amelogenesis Imperfecta – UK - 2011

The Most Published Target Enrichment Platform

>1000 Publications...

Nucleic Acids Research Advance Access published June 29, 2010

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh^a, Ming K. Lee^a, Silvia Casadei^a, Anne M. Thornton^a, Sunday M. Stray^a, Christopher Pennil^b, Alex S. Nord^a, Jessica B. Mandell^a, Elizabeth M. Swisher^b, and Mary-Claire King^{a,1}

Katharine Hutchinson^a, Agnes Viale^a, Zhongming Zhao^a, Roman K. Thomas^{2,7,8} and William Pao^{5,*}

Jenny L. Xiang^a, Maria E. Furtado^a, Jose Delgado^a, Chantal Fournier^a, Bernard Bryn C. Schierberl^{1,2}, Michel Michaelides^{9,10,11}, Richard G. Weleber⁹ and Joseph J. Higgins^{1,*}

Michael R. Stratton^{1,15} & P. Andrew Futreal¹

Herrmann^a, Kurt Zatloukal^a, Hans Lehrach^a, Michal R. Schweiger^a

Frederick R. Obeidat^{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000}

... and many more to come



For Research Use Only.
Not for use in diagnostic procedures.

Most Databases Covered for Comprehensive Variant Detection-Exons (Exome, WES)

| Parameter | Clinical Research | | Translational Research | | Cancer Research |
|--|-------------------------|---------------|------------------------|-------------|--------------------|
| | Clinical Research Exome | Focused Exome | All Exon V6+UTR | All Exon V6 | All Exon V6+COSMIC |
| Capture Size | 54Mb | 12Mb | 79Mb | 58Mb | 64Mb |
| CCDS | X | | X | X | X |
| RefSeq | X | | X | X | X |
| GENCODE | X | | X | X | X |
| miRBase | X | | X | X | X |
| TCGA | X | | X | X | X |
| UCSC | X | | X | X | X |
| HGMD, OMIM, ClinVar | X | X | X* | X* | X* |
| COSMIC | | | | | X |
| Amount of Seq. (for 80% at 20x) | 4Gb | <1Gb | 6Gb | 4Gb | 4Gb |
| Overall Workflow | 7h-1.5d | 4.5 days | 7h-1.5d | 7h-1.5d | 7h-1.5d |
| Add Custom Content | Yes | Yes | No | Yes | No |

SureSelect – The Leader in Target Enrichment

COMPLETE Solution

Best **PERFORMANCE**

FLEXIBLE Designs

SureSelect Technology

Most Complete Enrichment Solution

**All Exon
Designs**

**Human All
Exon**

**Non Human
Exomes**

**Clinical Research
Exome**

**Custom
Solutions**

**Custom
DNA**

**Custom
RNA**



**Targeted
Panels**

**Inherited
Disease**

**Human Kinome
(DNA/RNA)**

**Human
Methyl-Seq**

**ClearSeq Comprehensive
Cancer**

Post-capture

**SureSelect ^{XT}
Illumina**

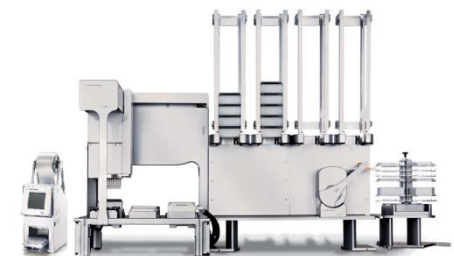
**SureSelect ^{0XT}
Illumina**

**SureSelect
Strand Specific
RNA
Illumina**

**SureSelect Ion
Torrent PGM**

Pre-capture

**SureSelect ^{XT2}
Illumina**



Automation



Agilent Technologies

For Research Use Only.
Not for use in diagnostic procedures.

Sureselect Library Preparation kits

| | CONSTITUTIONAL | | | | |
|-----------------------|--|---|--|---|---|
| | CANCER | | | | |
| | TARGET ENRICHMENT | | | | WGS |
| | SureSelect ^{XT} | SureSelect ^{XT} Fast | SureSelect ^{XT2} | SureSelect ^{QXT} | SureSelect ^{QXT} |
| Library Prep Features | <ul style="list-style-type: none"> • 3ug or 200ng • Best complexity | <ul style="list-style-type: none"> • 200ng input • High complexity • Tested with FFPE • 11hr sample to sequencing workflow | <ul style="list-style-type: none"> • 1ug or 100ng input • Streamlined: Precap Pooling + MM Reagents | <ul style="list-style-type: none"> • 50ng input • Single day sample to sequencing • Streamlined: transposase-based | <ul style="list-style-type: none"> • 50ng input • Single day sample to sequencing • Streamlined: transposase-based |
| Key Benefit | <ul style="list-style-type: none"> • Complexity for rare allele detection • Compatibility with low input | <ul style="list-style-type: none"> • Complexity for rare allele detection • Low input | <ul style="list-style-type: none"> • Cost effective | <ul style="list-style-type: none"> • Low input • One day workflow | <ul style="list-style-type: none"> • Low input • One day workflow |



SureSelect

FFPE Samples

FFPE Samples Using SureSelect

FFPE Challenges

- DNA quality issues (age, storage, fixation process)
- Sample degradation can impact sequencing quality (sensitivity, quality, allelic balance, etc)



FFPE Samples Using SureSelect

Questions

- Can FFPE-derived material be used for targeted sequencing?
- Is there any bias introduced by using FFPE vs fresh-frozen samples?
- How does target-enriched sequencing data derived from FFPE samples compare to microarray data?

RESEARCH ARTICLE

Open Access

Targeted high throughput sequencing in clinical cancer Settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity

Martin Kerick^{1†}, Melanie Isau^{1,2†}, Bernd Timmermann¹, Holger Sültmann³, Ralf Herwig¹, Sylvia Krobitch¹, Georg Schaefer^{4,5}, Irmgard Verdorfer^{5,6}, Georg Bartsch⁴, Helmut Klocker⁴, Hans Lehrach¹ and Michal R Schweiger^{1*}

Method

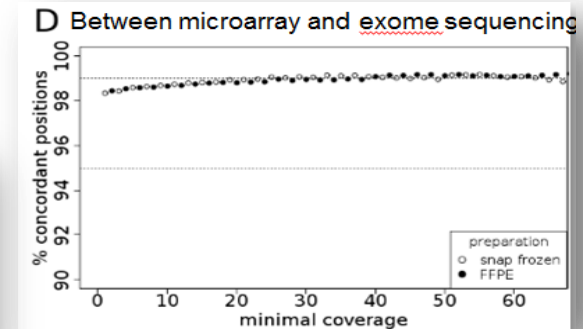
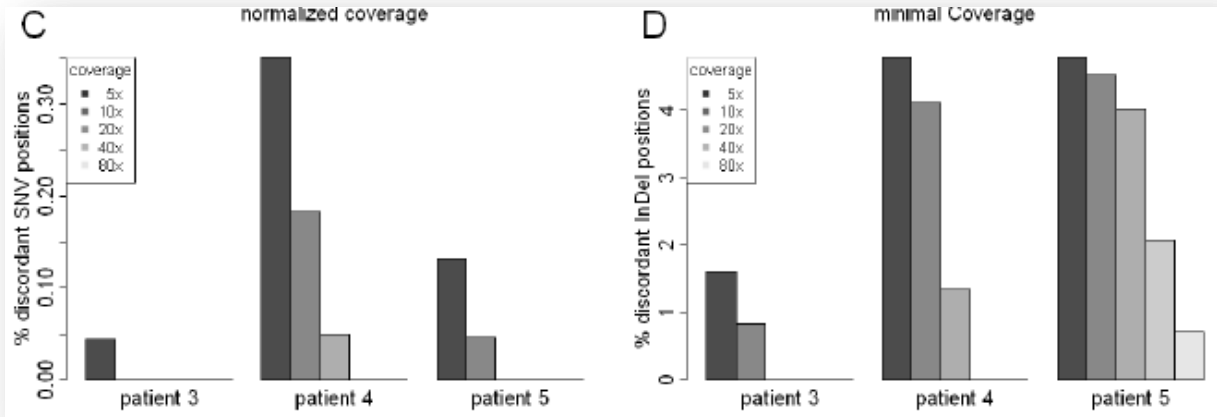
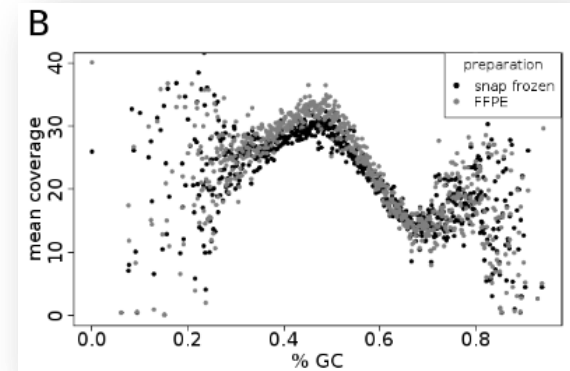
- 50Mb All Exon capture library and 3.9Mb custom capture
- 2 x 36bp single-end sequencing on Illumina GAIIx



FFPE Samples Using SureSelect

Results

- Similar % reads on target (75.5% frozen; 74.7% FFPE)
- Comparable GC-dependent coverage profile
- High NGS SNP concordance to microarray
- 98% concordance in SNP detection across varying input
- At 20X depth, <0.05% discordant SNPs between 2 foci for all patients



FFPE Samples Using SureSelect

Breakthrough in Breast Cancer Research

- Demonstration of FFPE samples from ICR with 50ng of input using SureSelect Human All Exon and modified protocol

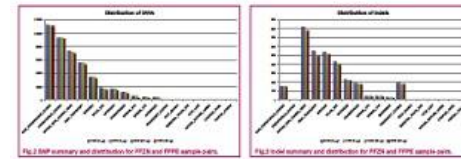
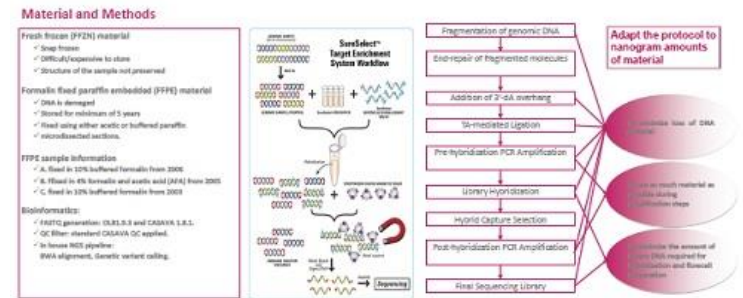
Major conclusions

- The 50ng method had comparable to 3µg for sequencing/enrichment metrics and variant calling
- The 50ng FFPE-based method was also comparable to 3µg protocol

ICR Exome re-sequencing from nanogram amounts of formalin-fixed paraffin embedded (FFPE) tumour material using the Agilent platform

The Institute of Cancer Research
 iwanka Kozarewa, Lina Chen, James Campbell, Kerry Fenwick, Ioannis Assiotis, Alan MacKay, Rachael Natrajan, Jorge Reis-Filho, Christopher J. Lord, and Alan Ashworth
 The Breakthrough Toby Robins Breast Cancer Research Centre, Institute of Cancer Research, London, UK

Introduction
 Formalin-fixed paraffin embedded (FFPE) treatment has been used for decades in clinical research since it maintains morphological features of the original tissue. The large archives of FFPE specimens that exist are a precious material source for cancer studies. However, all current exome re-sequencing platforms (provided by Illumina Inc., Agilent, and Nimblegen Roche) are restricted to microgram amounts of fresh frozen material (FFZN), which limits their application in tumour biology. This poster presents a series of modifications to the standard Agilent SureSelect protocol that allows for it to be used with nanogram amounts of FFPE material. The datasets obtained from FFPE-based libraries had a comparable sequencing and enrichment quality metrics to libraries prepared from microgram amounts of FFZN tumour material.



Conclusions

- The 50ng method had comparable to the 3µg one sequencing/enrichment metrics and variant calling.
- The 50ng FFPE based method was also comparable to the 3µg one.
- The only problem with the 50ng FFPE based method was the reduced (with 10-20%) percentage of unique reads which resulted in lower coverage. This could be rescued by increasing sequencing yield.

BREAKTHROUGH
 BREAST CANCER

SureSelect Custom DNA



The Best Performance and Flexibility

Capture only what you want

Custom content from <200Kb up to 24Mb+

Higher sample throughput per run

>1,000 samples/run with 96 indexes!

Cost effective

Minimize enrichment and sequencing costs

Faster time from sample to results

Greater sample throughput, easier data analysis

SureSelect – Proven High Sensitivity

Capture SNPs and Indels, while maintaining allelic balance



Frequent Mutations of Chromatin Remodeling Gene *ARID1A* in Ovarian Clear Cell Carcinoma

Siân Jones,¹ Tian-Li Wang,² Je-Ming Shih,³ Tsui-Lien Mao,⁴ Kentaro Nakayama,⁵ Richard Roden,³ Ruth Glas,⁶ Dennis Slamon,⁶ Luis A. Diaz Jr.,¹ Bert Vogelstein,¹ Kenneth W. Kinzler,^{1*} Victor E. Velculescu,^{1*} Nickolas Papadopoulos^{1*}



Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh^a, Ming K. Lee^a, Silvia Casadei^a, Anne M. Thornton^a, Sunday M. Stray^a, Christopher Pennil^b, Alex S. Nord^a, Jessica B. Mandell^b, Elizabeth M. Swisher^b, and Mary-Claire King^{a,1}

^aDepartments of Medicine and Genome Sciences and ^bObstetrics and Gynecology, University of Washington, Seattle, WA 98195

Table 1. Mutations in *ARID1A*, *KRAS*, *PIK3CA*, and *PPP2R1A* in human ovarian clear cell carcinomas.*

| Sample† | Gene | Transcript accession | Nucleotide (genomic)‡ | Nucleotide (cDNA)§ |
|---------|---------------|----------------------|--|---------------------------|
| OCC01PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26972561_26972562insA | c.3854_3855insA |
| OCC02PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26896034C>T | c.53C>T |
| OCC02PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26978879_26978880dupGT | c.5903_5904dupGT |
| | | | g.chr1:26972009_26972034delTGATGGGGCG | c.3659_3684delTGATGGGGCG |
| OCC03PT | <i>ARID1A</i> | CCDS285.1 | CATGTCCTATGAGCCA (hom) | CATGTCCTATGAGCCA |
| OCC07PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26896066C>A | c.585C>A |
| OCC08PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26970389delC | c.3391delC |
| OCC10PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26972790_26972792dupGCA (hom) | c.4001_4002dupGCA (hom) |
| OCC10PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26979804_26979805delTG (hom) | c.6827_6828delTG (hom) |
| OCC11PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26930334_26930335insCCTAC | c.145 |
| OCC13PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26974233_26974234insTGGC | c.492 |
| OCC14PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26972886_26972887_delIT (hom) | c.401 |
| OCC15PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26973940G>A | c.4635G>A |
| OCC15PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26978178T>A | c.5202T>A |
| OCC16PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26895967_26895973delCGCCGCC (hom) | c.486_492delCGCCGCC (hom) |
| OCC18PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26971925delA | c.3575delA |
| OCC20PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26970221delG | c.3223delG |
| OCC22PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26979694dupG | c.6718dupG |
| OCC23PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26896379_2689637980_insCGTC | c.898_899insC |
| OCC23PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26979686_26979687insT | c.6710_6711insT |
| OCC24PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26930542C>T | c.1663C>T |
| OCC27PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26896263_26896272delCGTCGTCTTC | c.782_791delCGTCGTCTTC |
| OCC27PT | <i>ARID1A</i> | CCDS285.1 | g.chr1:26971984_26971994delCAGCCCAGTAT | c.3634_3644delCAGCCCAGTAT |

25bp deletion

7bp deletion

10bp deletion

11bp deletion

Table 3. Genomic deletions and duplication identified by the assay

| Gene | Genomic event | Chromosome | Mutant sites identified by assay | | | |
|--------------|----------------------|------------|----------------------------------|------------|-----------|--------|
| | | | Start* | End* | Size (bp) | Ratio† |
| <i>BRCA1</i> | Deletion exons 1–15 | 17 | 41,226,145 | 41,327,157 | 101,013 | 0.509 |
| <i>BRCA1</i> | Duplication exon 13 | 17 | 41,230,562 | 41,235,836 | 5,275 | 1.578 |
| <i>BRCA1</i> | Deletion exons 14–20 | 17 | 41,203,975 | 41,229,297 | 25,323 | 0.519 |
| <i>BRCA1</i> | Deletion exon 17 | 17 | 41,219,596 | 41,219,755 | 160 | 0.495 |
| <i>BRCA2</i> | Deletion exons 1–2 | 13 | 32,889,020 | 32,890,900 | 1,881 | 0.489 |
| <i>BRCA2</i> | Deletion exon 21 | 13 | 32,950,734 | 32,952,070 | 1,337 | 0.544 |

*Breakpoints are flanked by *Alu* and other repeats, which are not captured.
†Reads per base pair for deletion or duplication/reads per base pair for wild-type genotype.

Large deletions and duplications



For Research Use Only.
Not for use in diagnostic procedures.

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh^a, Ming K. Lee^a, Silvia Casadei^a, Anne M. Thornton^a, Sunday M. Stray^a, Christopher Pennil^b, Alex S. Nord^a, Jessica B. Mandell^a, Elizabeth M. Swisher^b, and Mary-Claire King^{a,1}

^aDepartments of Medicine and Genome Sciences and ^bObstetrics and Gynecology, University of Washington, Seattle, WA 98195

- Inherited mutations in the tumor suppressor genes BRCA1, BRCA2, and multiple other genes predispose to high risks of breast and/or ovarian cancer.
- Developed an **Custom Panel** to detect all mutations in **21 genes** that predispose to breast or **ovarian cancer**, including BRCA1 and BRCA2
- There were **zero false-positive calls**.

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh^a, Ming K. Lee^a, Silvia Casadei^a, Anne M. Thornton^a, Sunday M. Stray^a, Christopher Pennil^b, Alex S. Nord^a, Jessica B. Mandell^a, Elizabeth M. Swisher^b, and Mary-Claire King^{a,1}

^aDepartments of Medicine and Genome Sciences and ^bObstetrics and Gynecology, University of Washington, Seattle, WA 98195

Table 2. Point mutations and small insertions and deletions identified by the assay

Excellent allelic balance

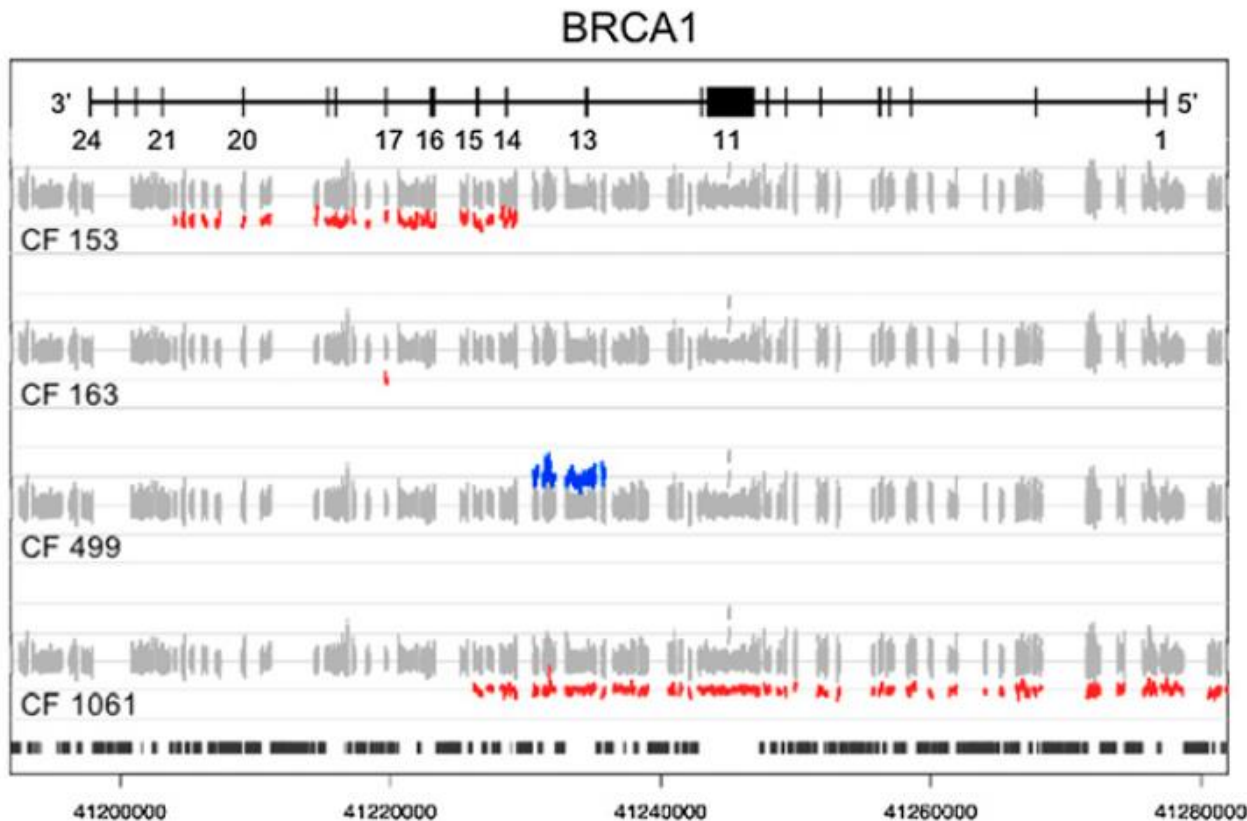
Deletion up to 19bp

| Gene | Nucleotide | Effect | Type | Size (bp) | Mutant sites identified | | | No. of reads | | |
|--------------|-----------------|-----------|--------------------|-----------|-------------------------|------------|------------|--------------|---------|-----------|
| | | | | | Chromosome | Start | End | Wild type | Variant | % Variant |
| <i>BRCA1</i> | 4510 del3ins2 | 1465 stop | Deletion-insertion | 1 | 17 | 41,228,596 | 41,228,597 | 525 | 596 | 0.53 |
| <i>BRCA1</i> | 5083 del19 | 1657 stop | Deletion | 19 | 17 | 41,222,949 | 41,222,968 | 700 | 644 | 0.48 |
| <i>BRCA1</i> | 5382 insC | 1829 stop | Insertion | 1 | 17 | 41,209,080 | 41,209,081 | 606 | 596 | 0.50 |
| <i>BRCA2</i> | 999 del5 | 273 stop | Deletion | 5 | 13 | 32,905,141 | 32,905,146 | 363 | 229 | 0.39 |
| <i>BRCA2</i> | 1983 del5 | 585 stop | Deletion | 5 | 13 | 32,907,366 | 32,907,371 | 304 | 258 | 0.46 |
| <i>BRCA2</i> | 6174 delT | 2003 stop | Deletion | 1 | 13 | 32,914,438 | 32,914,439 | 565 | 661 | 0.54 |
| <i>BRCA2</i> | 9179 C > G | 2984 stop | Nonsense | 1 | 13 | 32,953,650 | | 391 | 361 | 0.48 |
| <i>BRIP1</i> | 3401 delC | 1149 stop | Deletion | 1 | 17 | 59,761,006 | 59,761,007 | 651 | 486 | 0.43 |
| <i>CDH1</i> | 591 G > A | 157 stop | Nonsense | 1 | 16 | 68,842,406 | | 421 | 359 | 0.46 |
| <i>CHEK2</i> | 1100 delC | 381 stop | Deletion | 1 | 22 | 29,091,857 | 29,091,858 | 3,293 | 586 | 0.15 |
| <i>MLH1</i> | ivs14(-1) G > A | 568 stop | Splice | 1 | 3 | 37,083,758 | | 1,024 | 683 | 0.40 |
| <i>MSH2</i> | 1677 T > A | 537 stop | Nonsense | 1 | 2 | 47,693,895 | | 575 | 552 | 0.49 |
| <i>p53</i> | 721 G > A | R175H | Missense | 1 | 17 | 7,578,406 | | 449 | 306 | 0.41 |
| <i>PALB2</i> | 509 delGA | 183 stop | Deletion | 2 | 16 | 23,647,357 | 23,647,359 | 1,283 | 1,233 | 0.49 |
| <i>STK11</i> | ivs6(-1) G > A | 316 stop | Splice | 1 | 19 | 1,221,947 | | 722 | 572 | 0.44 |

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh^a, Ming K. Lee^a, Silvia Casadei^a, Anne M. Thornton^a, Sunday M. Stray^a, Christopher Pennil^b, Alex S. Nord^a, Jessica B. Mandell^a, Elizabeth M. Swisher^b, and Mary-Claire King^{a,1}

^aDepartments of Medicine and Genome Sciences and ^bObstetrics and Gynecology, University of Washington, Seattle, WA 98195



- CNVs can be measured by comparing read depth to a reference
- Ability to measure both deletions and duplications
- Variants range from 100s bp – +100kb

The Power of Custom Panels

3.19Mb

437 Custom Design
Mendelian Disorders

Detects fusion

Association of TALS Developmental Disorder with Defect in Minor Splicing Component *U4atac* snRNA

HUMAN GENOMICS

Carrier Testing
by Next-Gen

Callum J. Bell,^{1*} Darina
Elena E. Ganusova,¹

Faye D. Schilkey,¹ Vrunda Sheth,⁴ Jimmy E. Woodward,¹ Heather E. Peckham,⁴
Gary P. Schroth,³ Ryan W. Kim,¹ Stephen F. Kingsmore^{1,2†}

Patrick Edery,^{1,2*} Charles Marcaillou,^{3†} Mourad Sahbatou,^{4†} Audrey Labalme,^{1†}
Joelle Chastang,¹ Renaud Touraine,⁵ Emmanuel Tubacher,⁴ Faiza Senni,¹ Michael B. Bober,⁶
Sheela Nampoothiri,⁷ Pierre-Simon Jouk,^{8,9} Elisabeth Steichen,¹⁰ Siren Berland,^{11,12}
Annick Toutain,^{13,14} Carol A. Wise,¹⁵ Damien Sanlaville,^{1,2} Francis Rousseau,³
Françoise Clerget-Darpoux,^{16,17} Anne-Louise Leutenegger^{18,19}

seases

Ta
re

LETTE

D
si
of

Novel homo- and hemizygous mutations in *EZH2* in myeloid malignancies

Jullian Chmielecki,¹ Martin Peifer,¹ Peilin Jia,¹ Nicholas D. Socci,¹
Katherine Hutchinson,⁵ Agnes Viale,⁶ Zhongming Zhao,³
Roman K. Thomas,^{2,7,8} and William Pao^{5,*}

hearing

Tom Wason,¹ Ming Li,¹ Jia Li,¹ David G. Klapper,¹ Anne M. Thomson,¹ Sandy M. Gray,¹ Christopher Fennell,¹ Alex S. Nord,^a
Jessica B. Mandell,^a Elizabeth M. Swisher,^b and Mary-Claire King^{a,1}

A. Eliot Shearer^{a,b,1}, Adam P. DeLuca^{c,d,1}, Michael S. Hildebrand^{a,1}, Kyle R. Taylor^{c,d}, José Gurrola II^a, Steve Scherer^{e,2},
Todd E. Scheetz^{c,d,f,2}, and Richard J. H. Smith^{a,b,g,2,3}



Easily Create Your Custom Design

NEW!
Streamlined
Design Software



SureDesign
Free Web Portal



**Genomic
Coordinates**
Regions you want
to capture



Design Services
For Complex
Designs



**Your Custom
SureSelect Kit**
A bait library to capture
your regions of interest



Agilent Technologies

For Research Use Only.
Not for use in diagnostic procedures.

Custom Design Service Link

The screenshot shows the Agilent SureDesign web application. The browser address bar displays <https://earray.chem.agilent.com/suredesign/home.htm>. The page title is "SureDesign" and the current step is "Define Design".

Define Design Form:

- Design Name:
- Species: [Select ?](#)
- Build: H. sapiens, hg19, GRCh37, February 2009
- Create In: [Select](#)

Tooltip: Please contact Agilent Design Services if your species of interest is not available in this list.

Design Summary:

SureSelect DNA Design

| | |
|-----------|----------------|
| Name: | NA |
| Species: | NA |
| Category: | SureSelect DNA |

Target Regions

| | |
|------------|----|
| # Regions: | NA |
| Size: | NA |

Probes

| | |
|-------------|----|
| # Probes: | NA |
| Size: | NA |
| Price Tier: | NA |
| Coverage: | NA |

[UCSC View](#) [Download](#)

Custom SureSelect libraries are compatible with all Target Enrichment sample preparation kits and sequencing platforms supported by Agilent.

The Power of Custom Panels

For any species!



Neves et al. *BMC Proceedings* 2011, 5(Suppl 7):O48
<http://www.biomedcentral.com/1753-6561/5/S7/O48>



ORAL PRESENTATION

Open Access

Targeted sequencing in the loblolly pine (*Pinus taeda*) megagenome by exome capture

Leandro Neves^{1*}, John Davis², Brad Barbazuk³, Matias Kirst²

From IUFRO Tree Biotechnology Conference 2011: From Genomes to Integration and Delivery
Arraial d'Ajuda, Bahia, Brazil. 26 June - 2 July 2011

Saintenac et al. *Genome Biology* 2011, 12:R88
<http://genomebiology.com/2011/12/9/R88>



RESEARCH

Open Access

Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome

Cyrille Saintenac, Dayou Jiang and Eduard D Akhunov*





SureSelect Human Methyl- Seq for Illumina Sequencing

Agilent Gene Regulation
Solutions

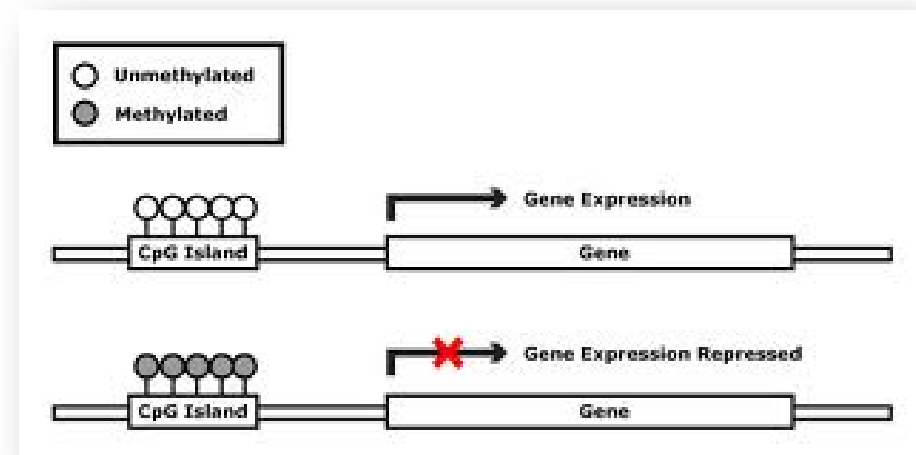
CpG Dinucleotides & Their Genomic Locations

CpG islands

- High frequency of CpG dinucleotides
 - ✓ $> 500\text{bp}$ & GC content $> 55\%$ & observed/expected CpG ratio > 0.65
- In or near about 40% of promoters of mammalian genes

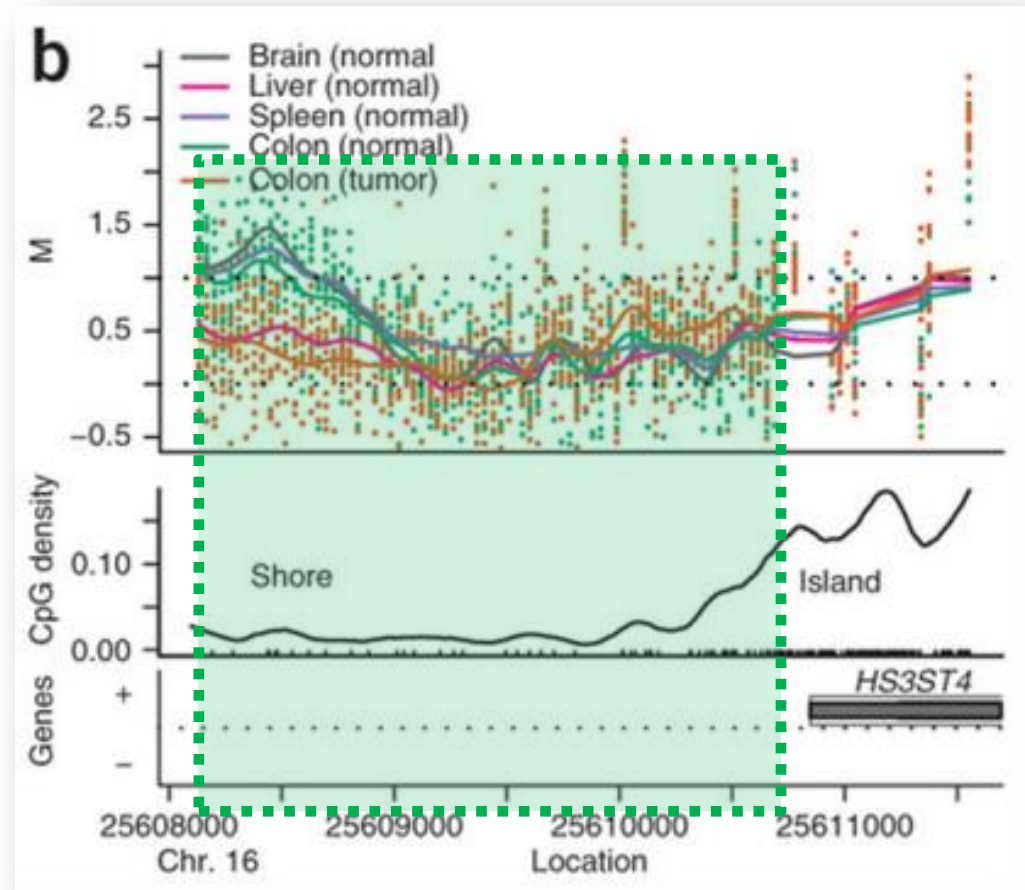
Promoters

- 75% of transcriptional start sites have CpG-rich regions
- 88% of active promoters are associated with CpG-rich sequences



Differentially Methylated Regions (DMR)

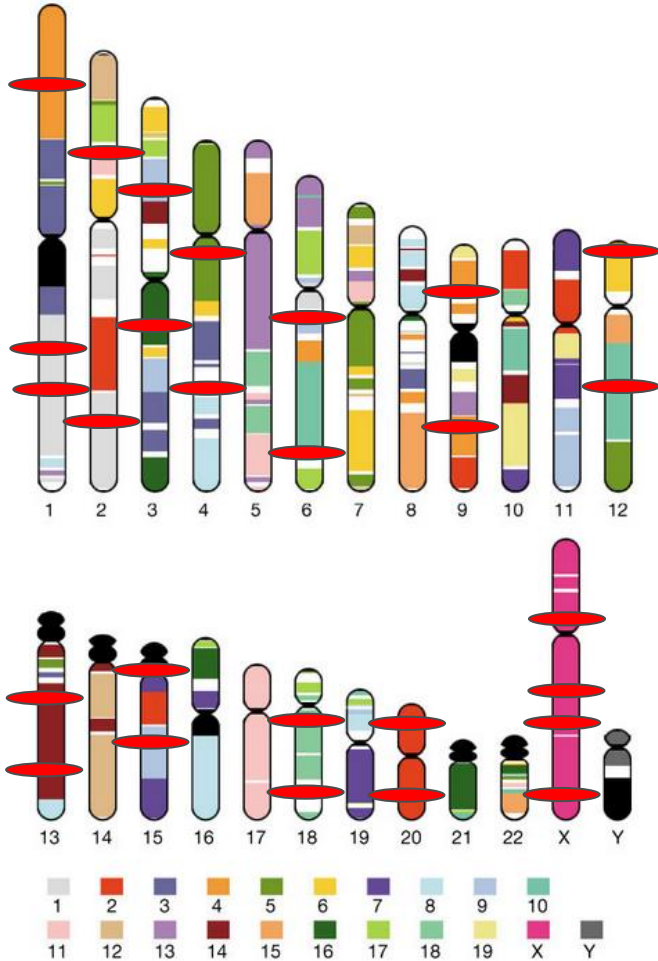
- CpG islands (e.g. 4~8 % tissue-specific differentially methylated regions or T-DMR)
- CpG island **shores** (~2kb away from islands, e.g. 76% of T-DMRs in shores)
- CpG island **shelves** (~4kb away from islands)



HS3ST4 : heparan sulfate D-glucosaminyl 3-O- sulfotransferase 4

Irizarry RA et al. Nature Genetics 2009

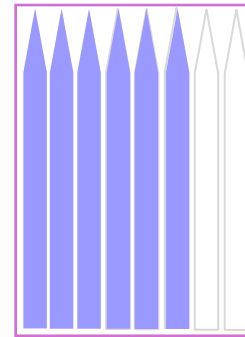
Target Known Methyl Regions



Less than 3% of the genome is reported to vary its methylation state

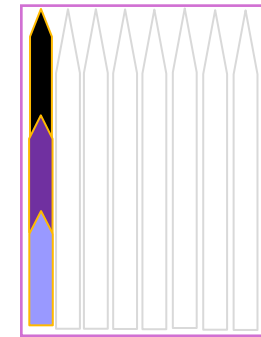
- Focus your research with SureSelect Methyl-Seq

Whole Genome Bisulfite Sequencing



3-6 lanes for one sample
90-180Gb per sample

SureSelect Human Methyl-Seq



3 samples per lane
10Gb per sample

9-18x more efficient with greater depth per region

SureSelect^{XT} Human Methyl-Seq

The First Comprehensive Methylation Discovery System

Discovery Tool

- Probes independent of Methylation state
- Determine methylation state of all methyl sites in region

Comprehensive Design

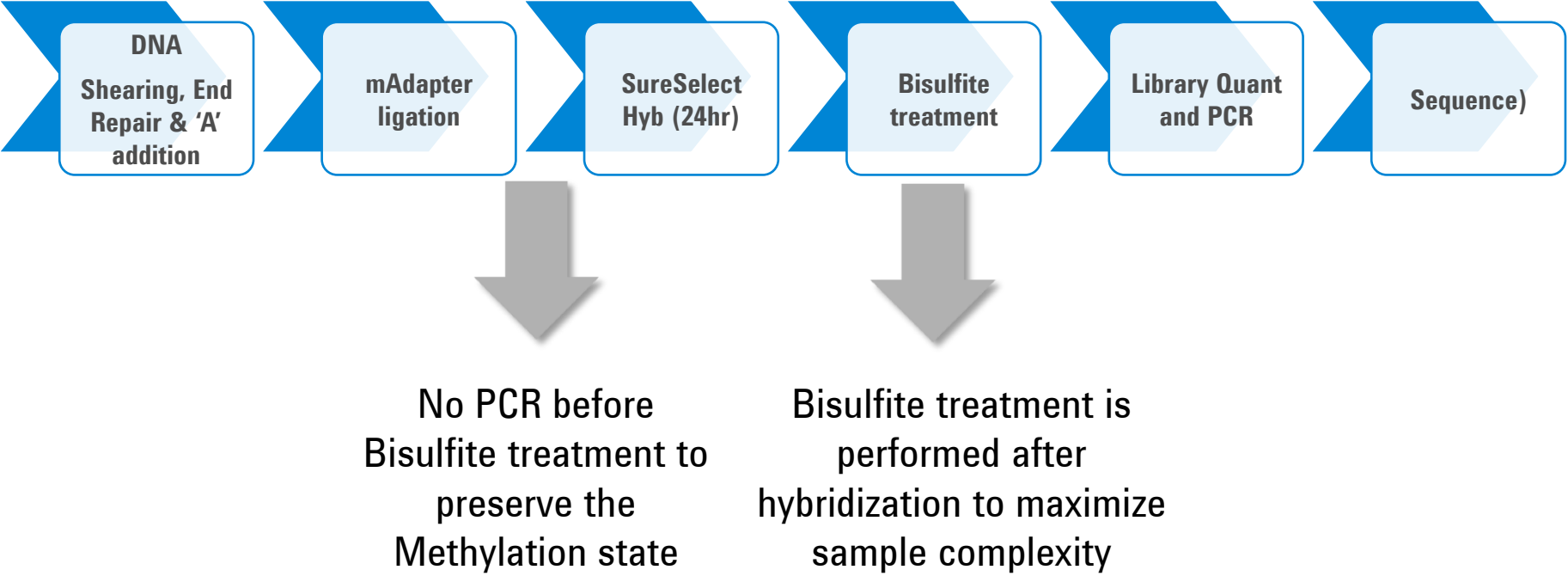
- CpG Islands
- Promoter regions
- DMRs (Differentially Methylated Regions)

DESIGN CONTENT - 84 Mb Design, 3.7M

CpGs

- CpG islands
- Cancer, Tissue-specific DMRs
- GENCODE promoters
- DMRs or regulatory features in:
 - ✓ CpG Islands, shores and shelves $\pm 4\text{kb}$
 - ✓ DNaseI hypersensitive sites
 - ✓ Refseq Genes
 - ✓ Ensembl Regulatory Features

SureSelect Methyl-Seq Protocol

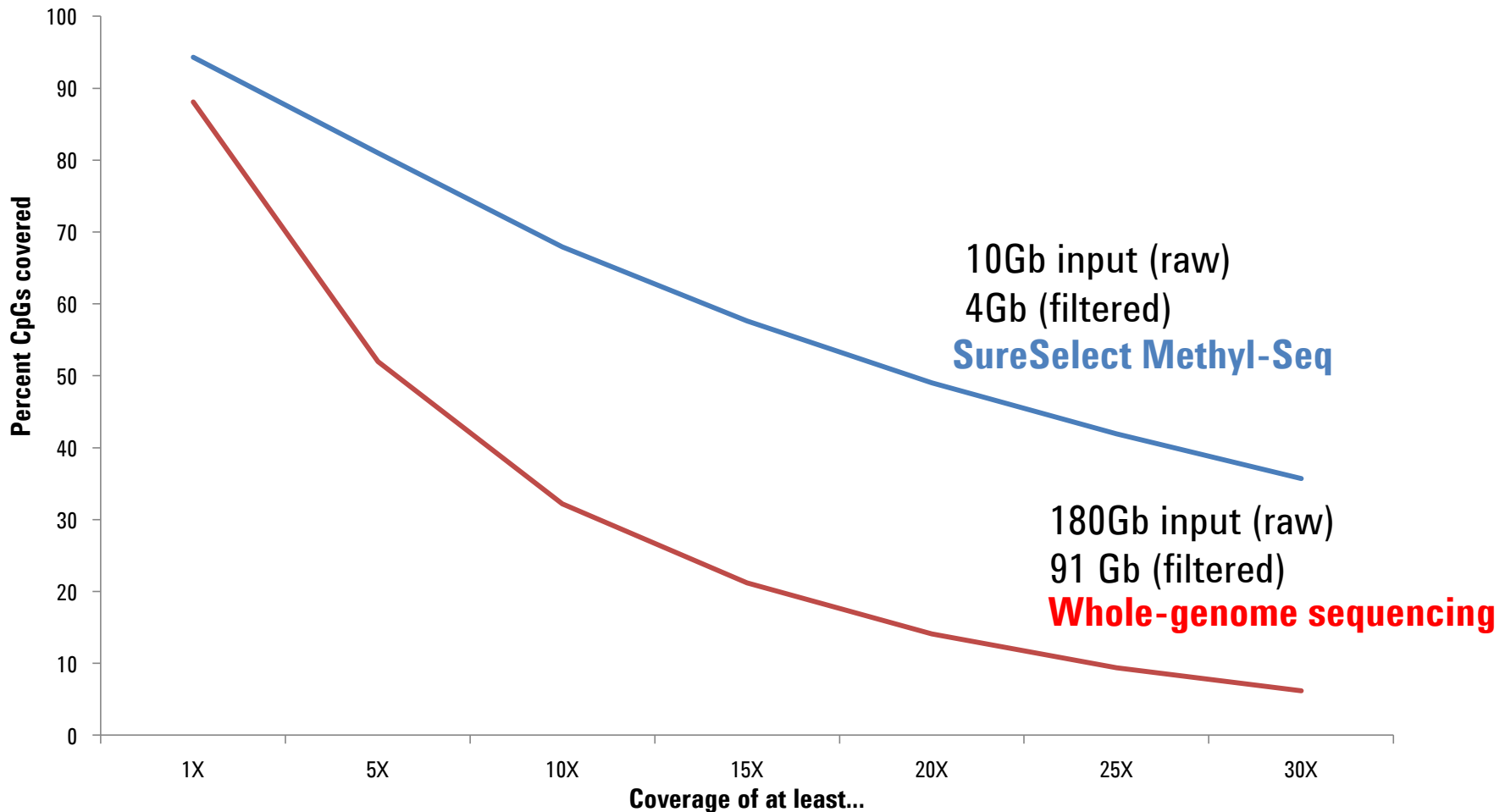


No PCR before
Bisulfite treatment to
preserve the
Methylation state

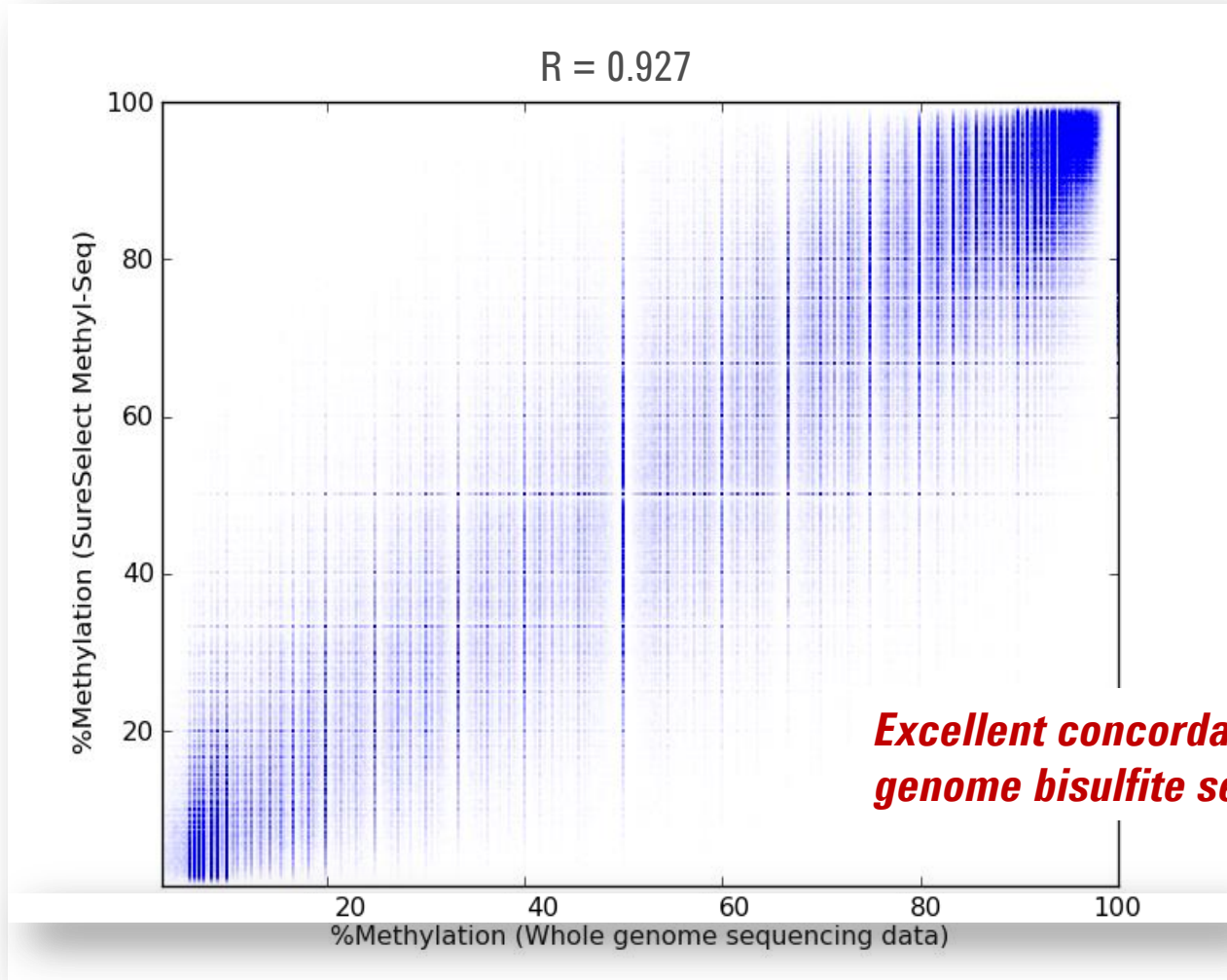
Bisulfite treatment is
performed after
hybridization to maximize
sample complexity

SureSelect vs. Whole-Genome Bisulfite Sequencing

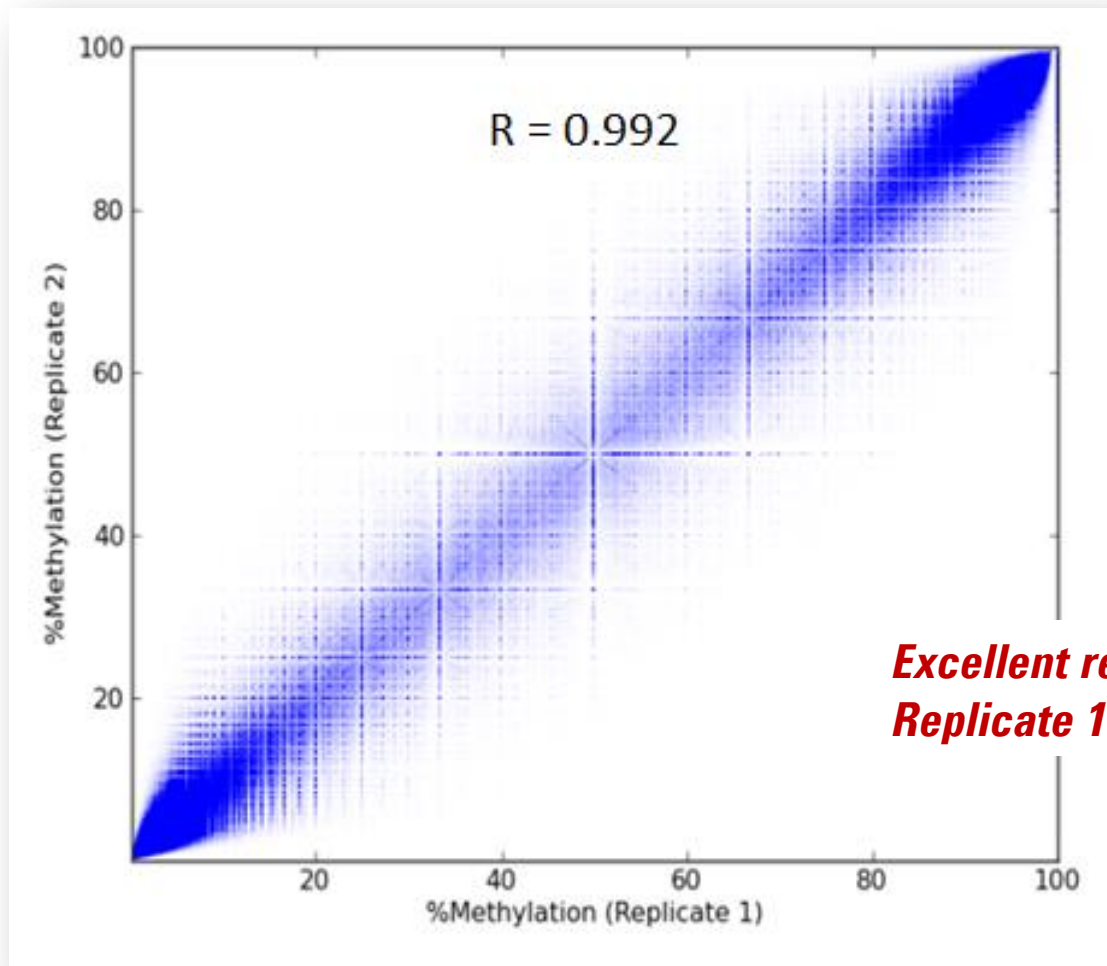
Coverage vs. % CpGs covered in targeted regions (~3.7 million CpGs)



Good Concordance between WGBS & SureSelect Methyl-Seq



Excellent Reproducibility

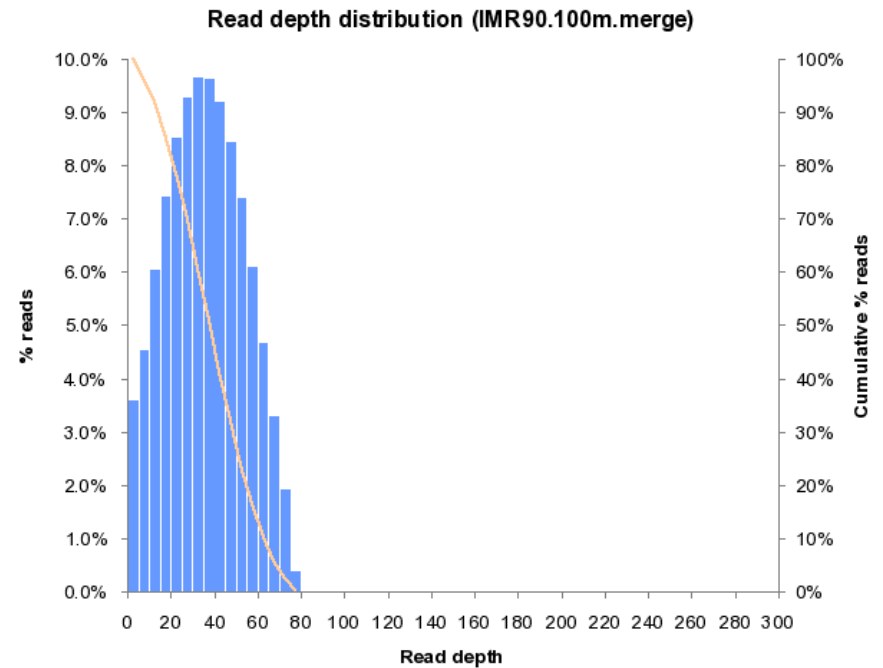


***Excellent reproducibility for IMR90
Replicate 1 vs. 2***

Capture performance

Input 10Gb Raw sequence

| | |
|---|------------|
| Percentage reads in targeted regions: | 82.0% |
| Percentage reads in regions +/- 100bp: | 93.6% |
| Percent of genome targeted: | 2.7% |
| Enrichment in targeted regions: | 30.07 |
| Uniformity (3/4 mean with upper tail): | 91.4% |
| Number of bases in targeted regions: | 84,367,621 |
| Percentage of targeted bases covered by... | |
| ...at least 1 read: | 98.7% |
| ...at least 10 reads: | 91.4% |
| ...at least 20 reads: | 78.9% |



Tissue Specific DMRs



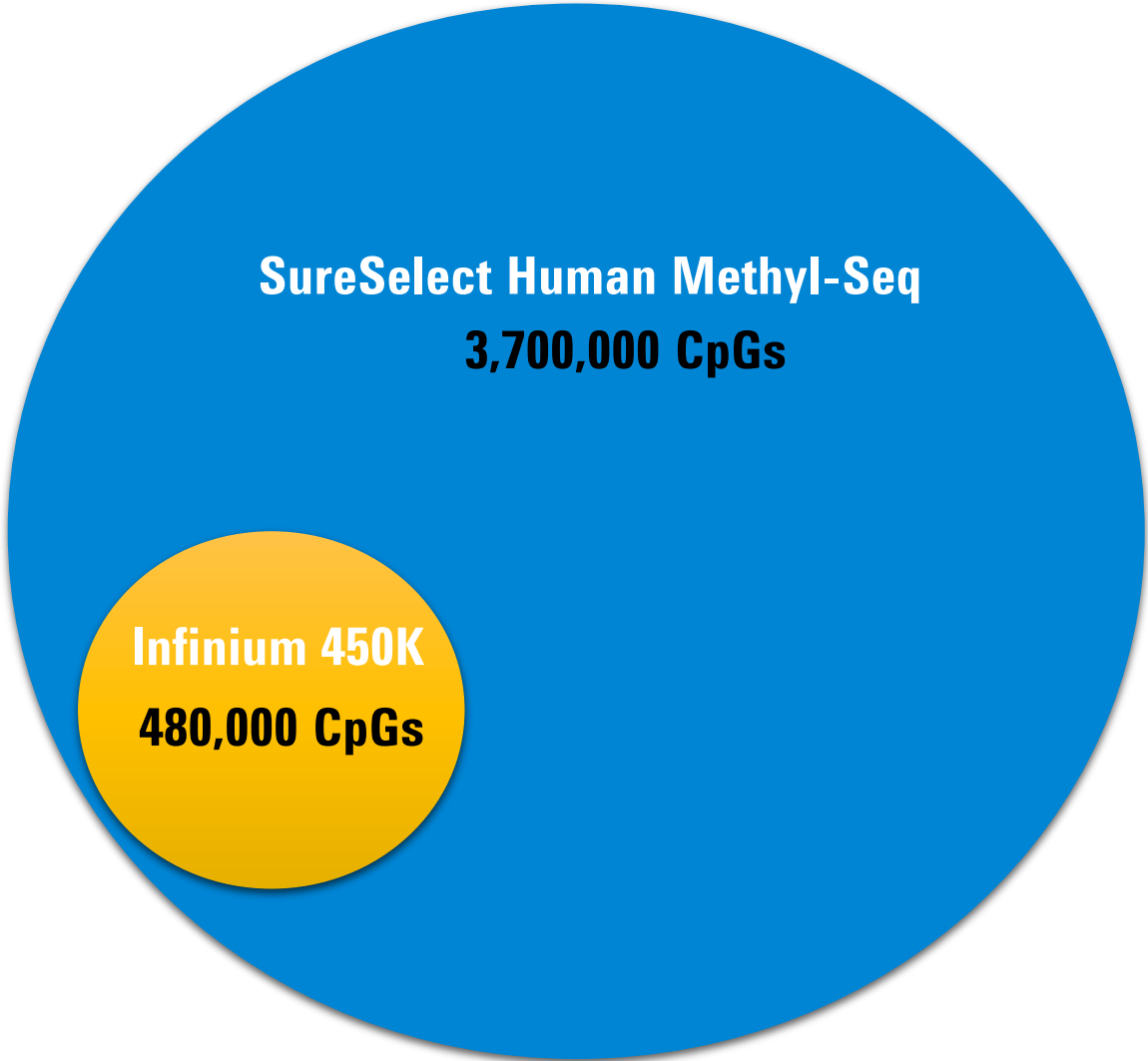
MeDip-Seq & Reduced Representation Bisulfite Sequencing (RRBS)



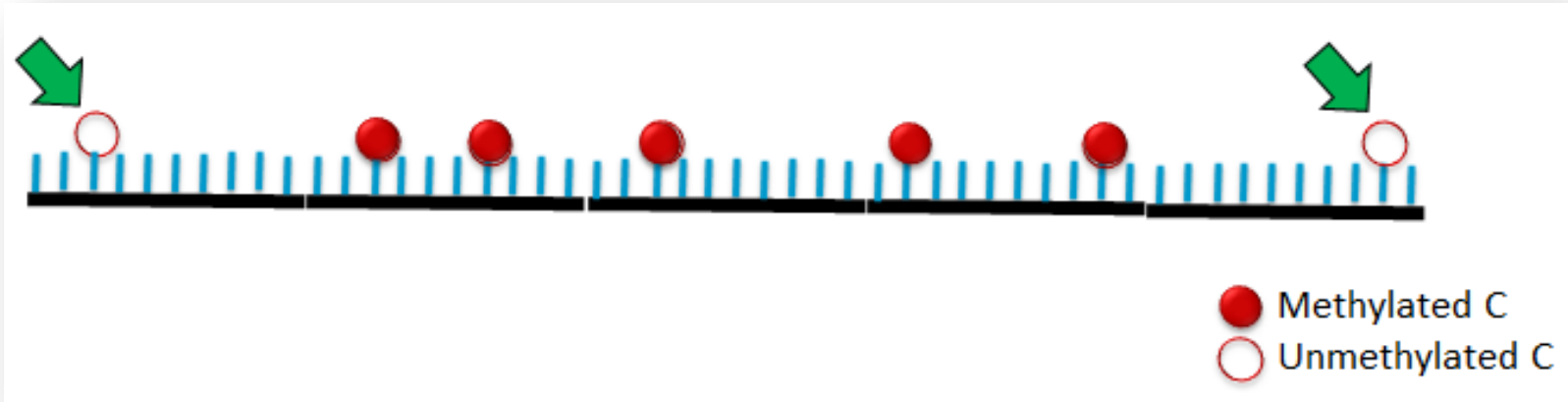
Limitations

- Cannot target specific regions (i.e. DMRs in Shelf and Shore regions)
- Biased towards methylated regions, Repeat sequences & CpG-rich sequences
- Can miss under-methylated regions
- Difficult to design since knowledge of methylation state for the target region is needed

More Coverage of CpGs Compared to Infinium 450K Array



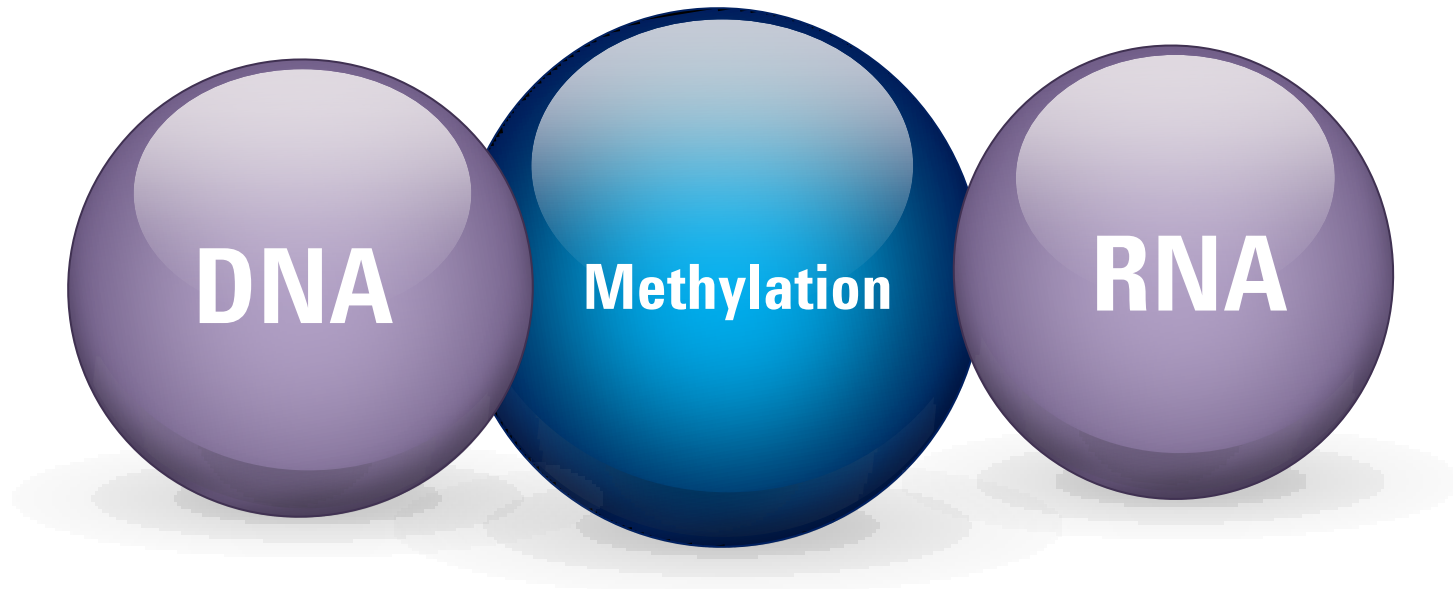
Microarray-Based Method: Infinium 450K Array



- Improves costs and throughput over whole-genome bisulfite
- Does not report individual methylation status
- Not whole-genome
- Does not cover 'shores and shelves' known to be important in DMR
- What are you missing?

SureSelect – “Omics” Solution

- **DNA:** Genetic variation
- **RNA:** Gene Expression
- **Methylation:** Effects on Gene Expression



HaloPlex^{HS}

Get to Know Your DNA. Every Single Fragment.



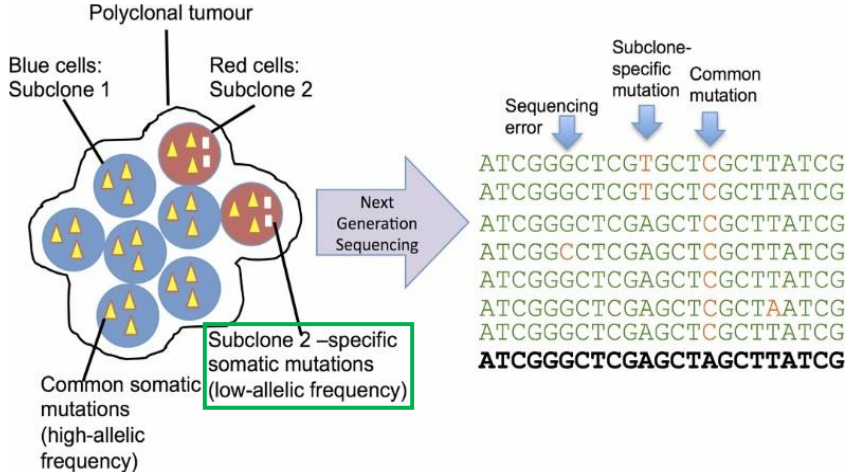
Low allele frequency variants

What are low allele frequency variants?

- Variants present at a frequency below 3%

What are low allele frequency variants implicated in?

- Clonal evolution and pathogenesis
- Tumor subclonal heterogeneity
- Immunological diversity



Adapted from Stead *et al* (2013) Human Mutation 34: 1432-1438

Low allele frequency variants

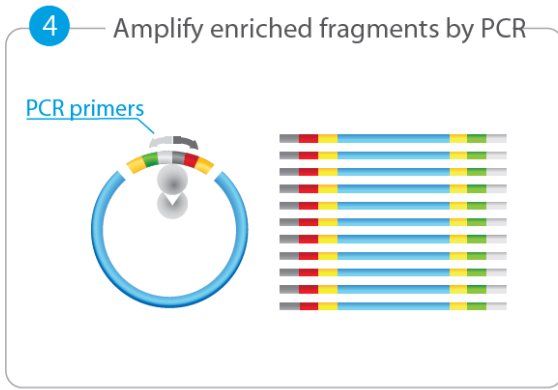
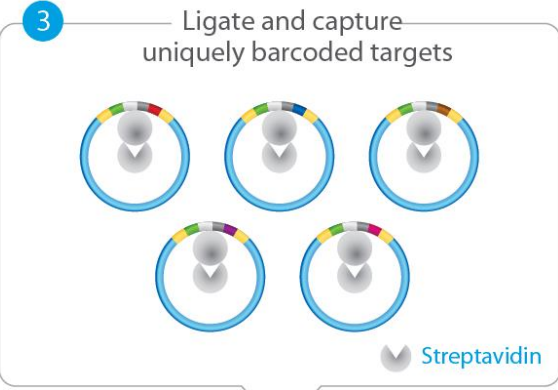
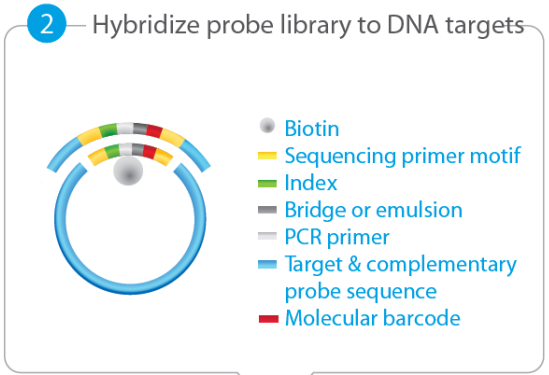
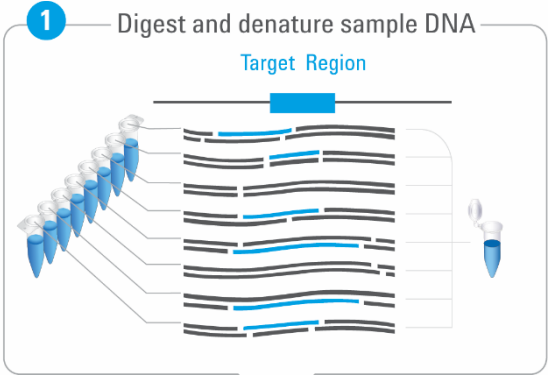
- Low allele frequency variants are difficult to detect by conventional NGS methods
- Relatively high error rate of sequencers (1 wrong base call in 100-1000 sequenced bases)

| Platform | Primary error type | Background (%) |
|---------------------|-------------------------------|----------------|
| Pacific Biosciences | G/C deletions | 16 |
| Life Ion Torrent | Short deletions, homopolymers | 1 |
| ABI SOLiD | A-T bias | 0.2 |
| Illumina MiSeq | Single nucleotide | 0.1 |
| Illumina HiSeq | Single nucleotide | 0.1 |

Kennedy *et al* (2014) Nature Protocols 9: 2586 - 2606

Requires molecular barcodes for increased sensitivity and accuracy

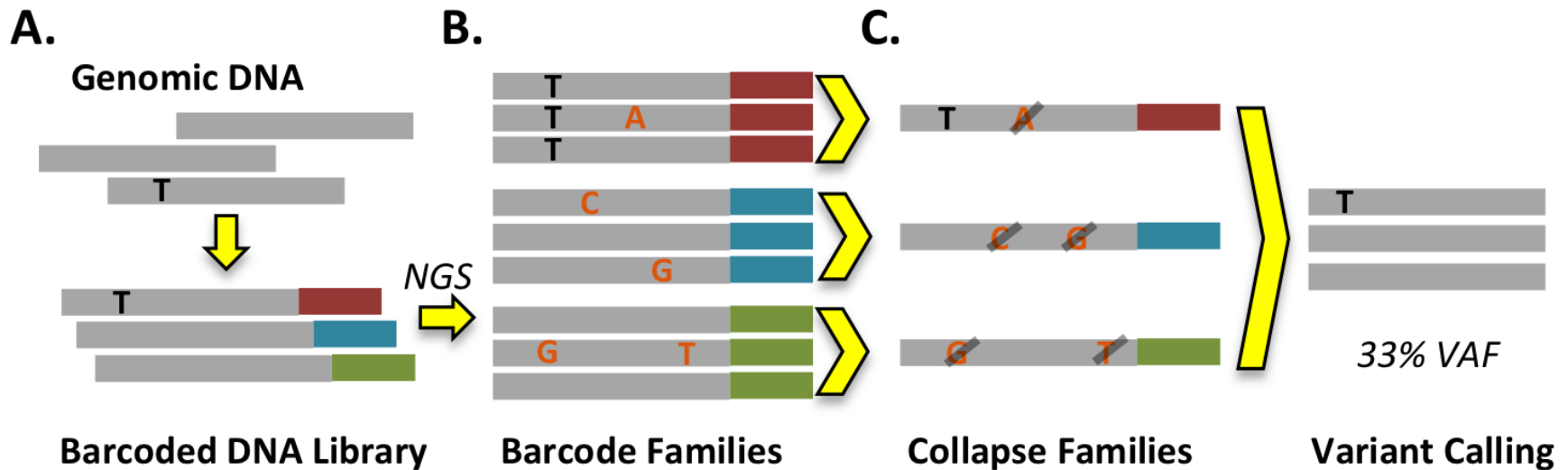
How HaloPlex^{HS} works



Molecular Barcodes

Molecular barcodes are degenerate oligonucleotide sequences (10-16bp) attached to individual DNA molecules

Allow for accounting of sequencer and PCR errors in high coverage NGS data



Courtesy of Dr. Eric Duncavage (AGBT 2015)

The need for sensitivity and accuracy

Performance of Common Analysis Methods for Detecting Low-Frequency Single Nucleotide Variants in Targeted Next-Generation Sequence Data

David H. Spencer,*
Eric J. Duncavage*

Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations

Glenn K. Fu,¹
and Stephen P.

Detection and quantification of rare mutations with massively parallel sequencing

Isaac Kinde, Jian Wu, Nick Papa

Detection of ultra-rare mutations by next-generation sequencing

Michael W. Schmitt¹

Detecting ultralow-frequency mutations by Duplex Sequencing

Scott R Kennedy¹
Marc J Prindle^{1,1}

Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation

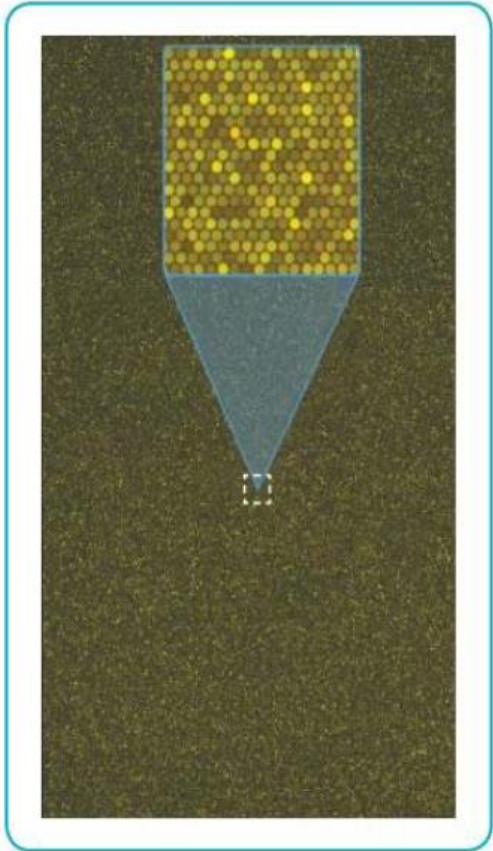
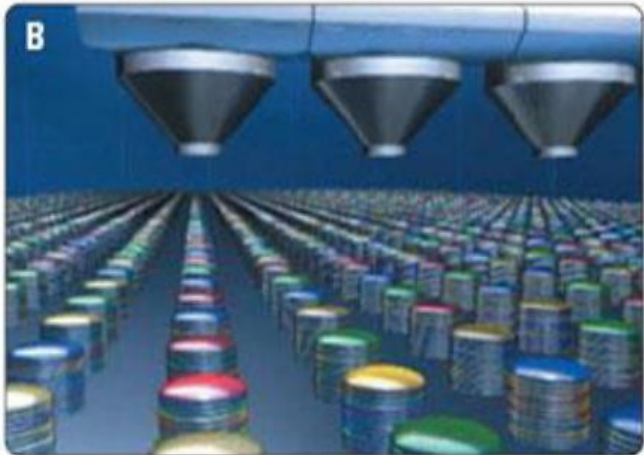
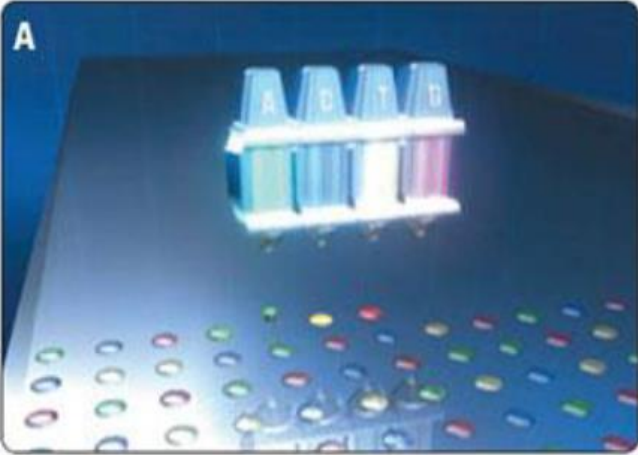
Joseph B. Hiatt, Colin C. Pritchard, Stephen J. Salipante, et al.





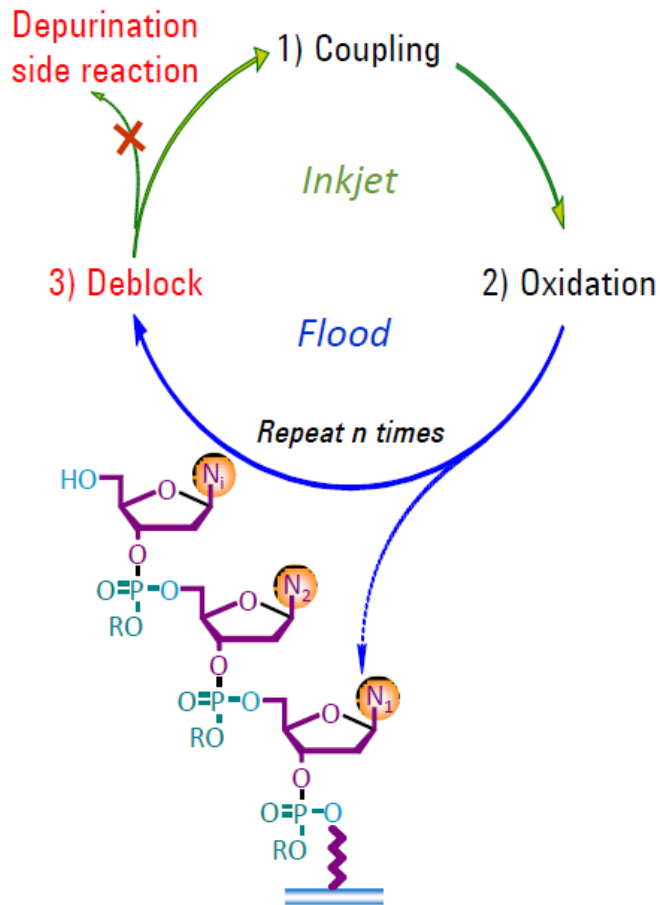
Oligo Library Access Program

SurePrint Technology



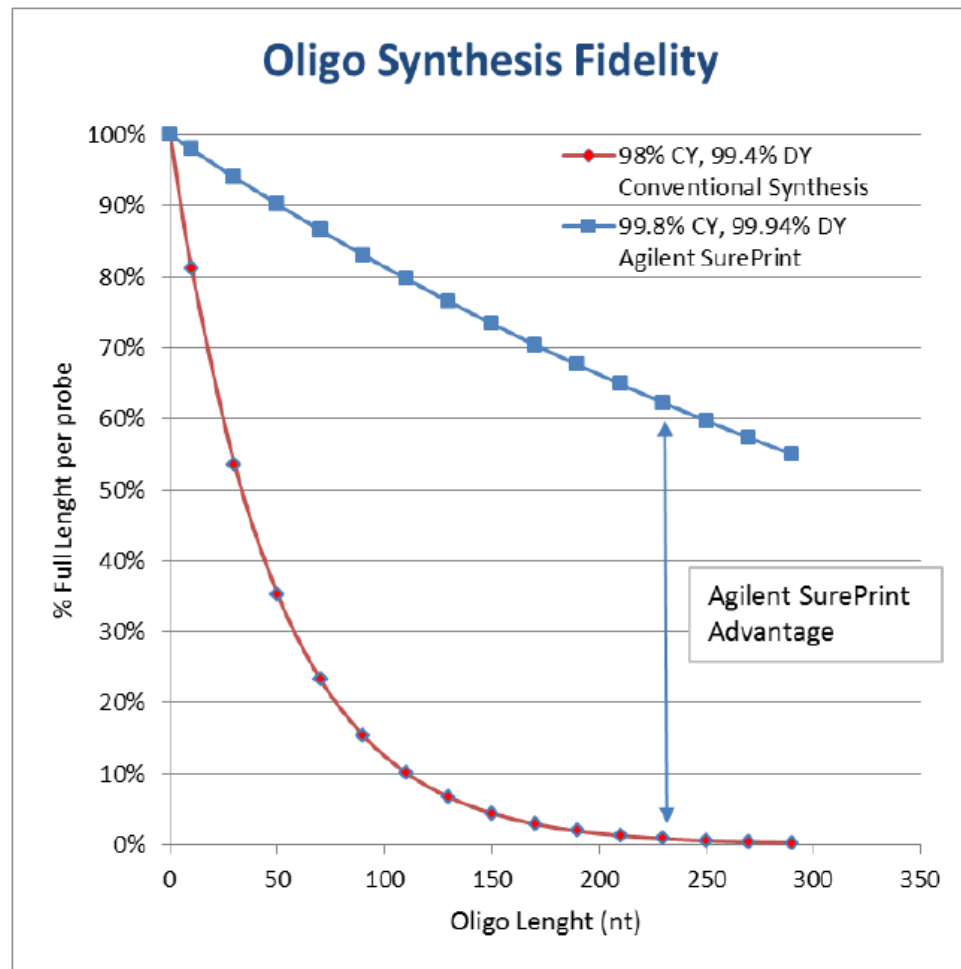
Agilent's million-featured microarray.
Inset shows a magnified view of a small region.

Chemical Synthesis: Achieving High Sequence Fidelity



Long length synthesis is achieved by improved cycle yield

- ↑ Coupling efficiency
- ↓ Depurination
- ↑ Consistency



$$\%FL = (CY * DY)^{nt}$$

%FL= %Full Length

CY=Synthesis Cycle Yield

DY=Depurination Cycle Yield

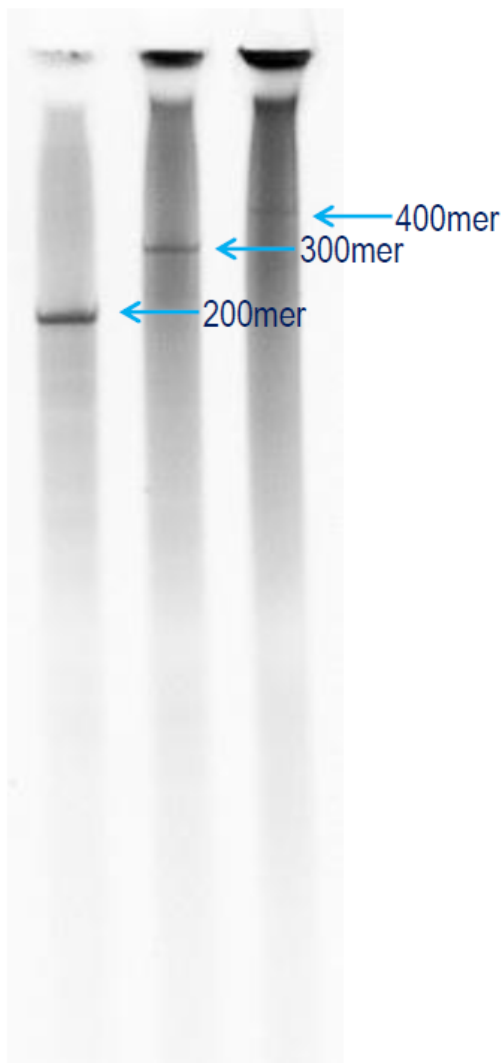


Agilent Technologies

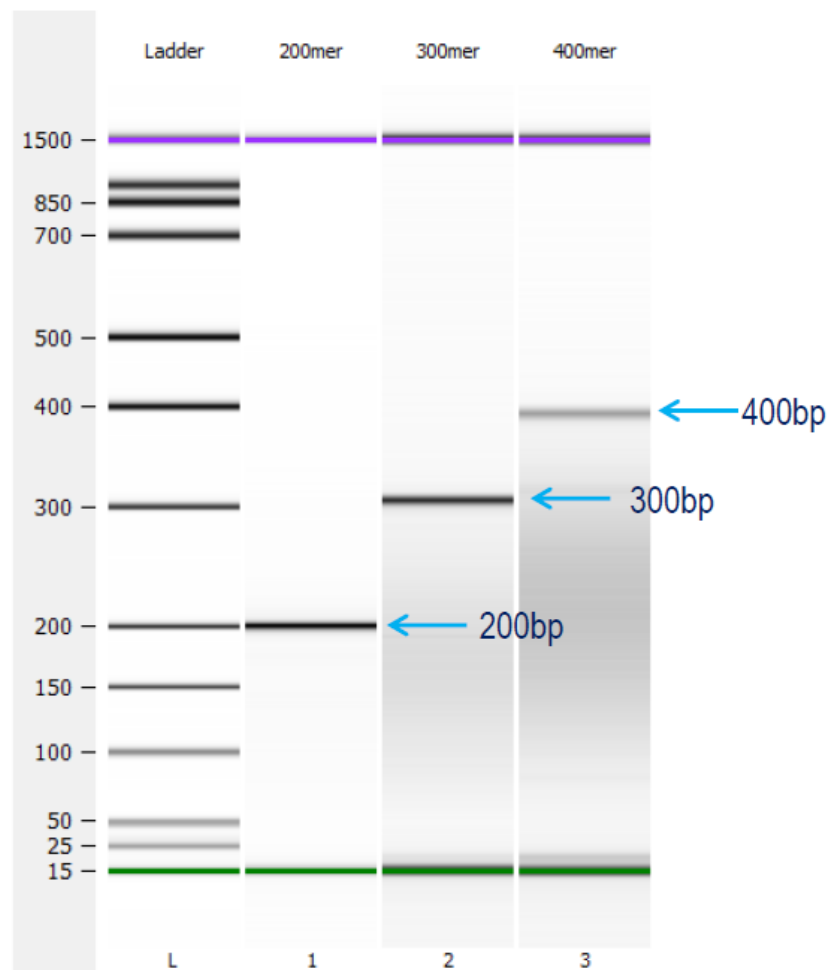
For Research Use Only.
Not for use in diagnostic procedures.

Pushing the limits of Chemical Synthesis in Research

OLS on denaturing gel



PCR of OLS up to 400mer

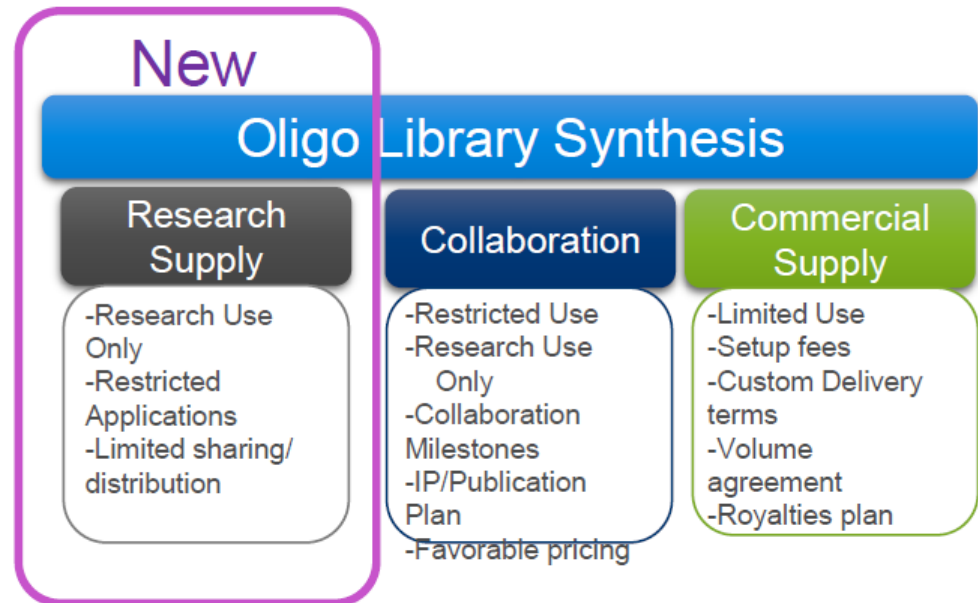


Oligo Library Synthesis Program (New Program)

Three paths to Agilent Custom Oligo libraries:

- **Research Supply:** Enable research applications using Agilent Oligo Library Synthesis products
- **Collaboration agreement:** Partner with Agilent in developing Novel Oligo libraries applications, typically leading to publication/Patent filling.
- **Commercial Supply:** Partner sees value on a continuous supply of Agilent Oligo libraries with negotiated use terms and agrees to a minimum annual volume. An upfront fee covers the setup of the products and agreement.

Common Terms: Partner should verify that it has obtained all third party required licenses for the given applications that might be covered under third party Intellectual Property.



Oligo Library Technology Access Program: Research Supply

Enable research applications using
Agilent Oligo Library Synthesis products

In scope for this program:

- Research Use Only (contact us for other uses)
- Oligo libraries up to 230nt long of complexities of 7500, 15000, 100000 and 244000 unique oligos/library on amounts of approx. 10pmols/library
- 4-5 week delivery time
- Competitive pricing
- Outside of Restricted Applications (see below), no IP / publication restrictions by Agilent. Customer is responsible for any third party licenses that might be needed for the intended use/application. Agilent will be indemnified by customer in case of infringements claims.

Restrictions:

- Libraries cannot be resold/distributed outside of purchasing institution
- Purchase and Use Restricted to applications NOT involving the following fields*:
 - Target enrichment/capture
 - In Situ Hybridization (ISH) and variants
 - Gene Assembly (assemble of 2 or more oligos into a longer DNA construct)
 - Genome editing applications, including Functional Genomics screening libraries
 - Site Directed Mutagenesis

*For use in Restricted Applications contact Agilent for alternative Products or Programs

SurePrint Oligo Library Supply

- **Content:**

- Library Complexities (Unique Sequences): **7.5k, 15k, 100K, 244K**
- Length: **20-230nt**. Sequences can have different length.
- Amount: **10pmol** (total library amount)
- Delivery Format: Library is supplied dried out in a single tube
- Pricing starting at 90bases
 - **Academic research as low as 0.04¢/base (\$0.4/kb)**. Additional sub-library charges may apply.

- **Library Design:**

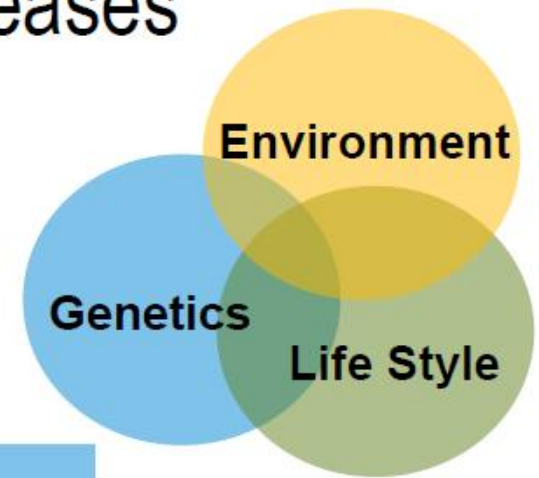
- Free design content, as long as not in conflict with restricted applications
- It can include priming regions to amplify the library, additional charge for multiple sub libraries

CRISPR Libraries: Early Access

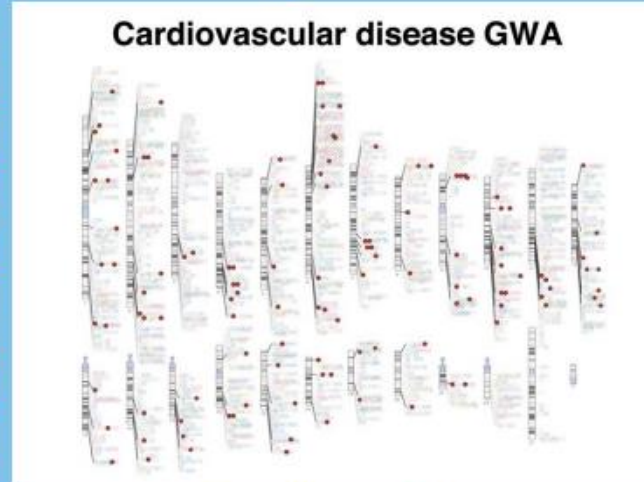
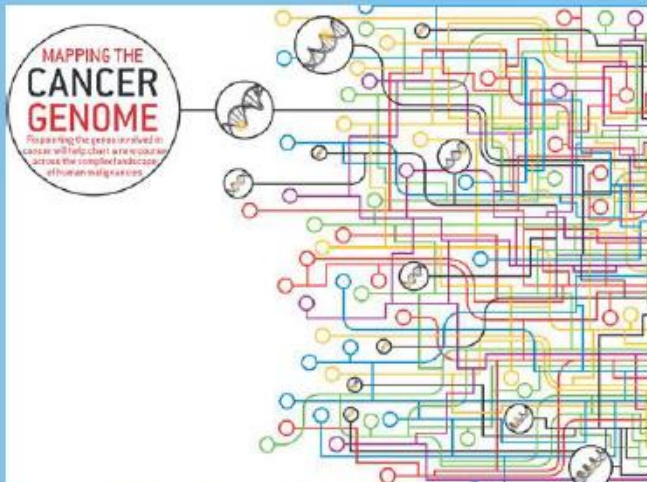


Life Sciences Research: Complex Diseases

- ❑ Cancer
- ❑ Cardiovascular disease
- ❑ Neurological disorders
- ❑ Metabolic disorders
- ❑ Autoimmune disorders



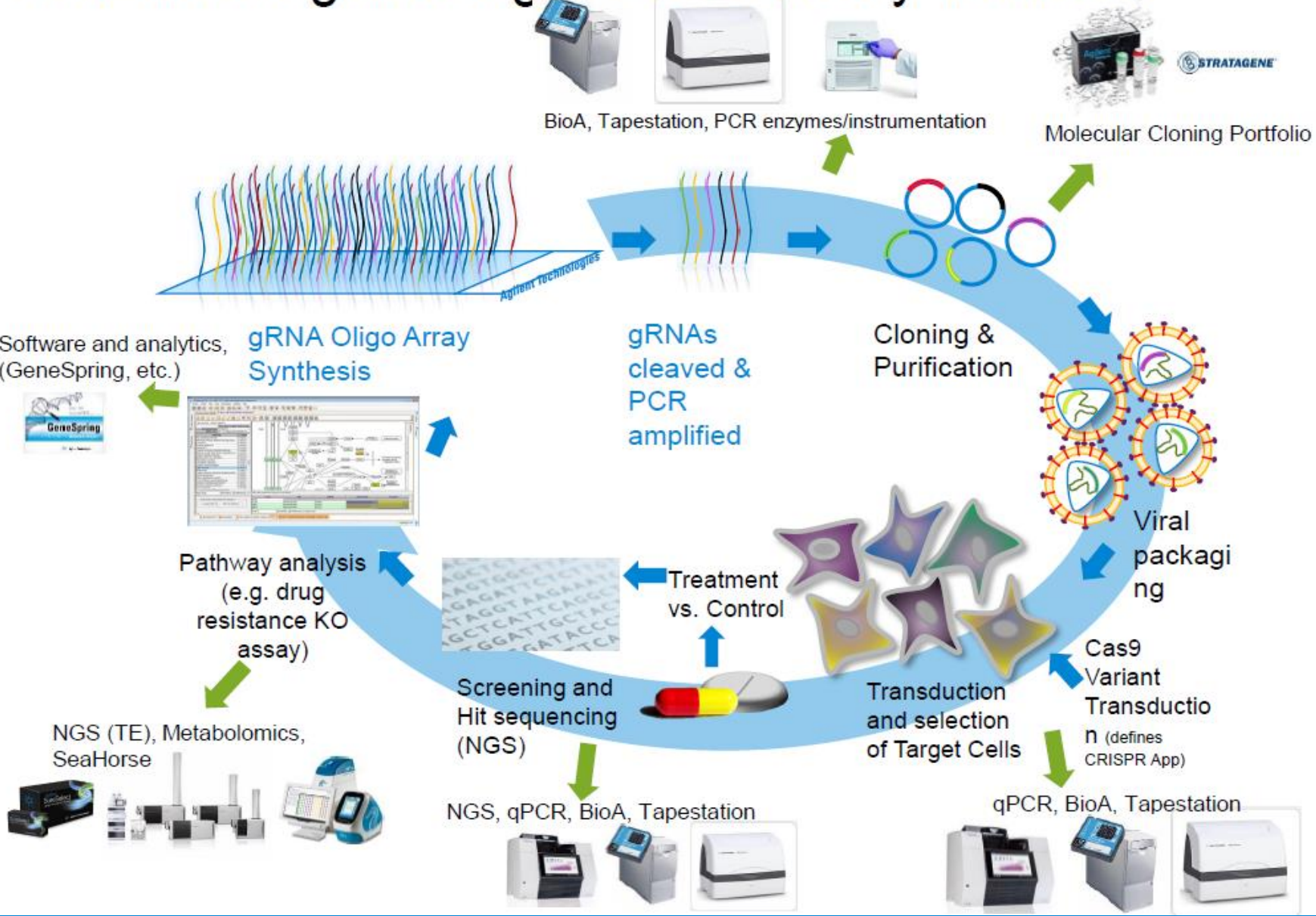
Genome sequencing has yielded 1000s of disease associated candidate genes
















Understanding the functions and roles of these genes has been challenging



Genome Engineering: Pooled library workflow



SureGuide CRISPR libraries

| | When (Early Access) | Usage | Applica tion | Strengths |
|---|------------------------|---------------------------------------|---|----------------------------------|
|  Catalog hExome | Now | Functional genomics/target ID | Knock-out libraries  | Highest Quality, Good Value |
|  Custom hExome | Now | Functional genomics/target validation | Knock-out subsets  | Highest Quality / Custom Designs |
|  Custom hGenomeWide | Now | Genome regulatory networks | CRISPR i/a  | Highest Quality / Custom Designs |
|  Catalog mExome | Now | Functional genomics/target ID | Knock-out libraries  | Highest Quality / Custom Designs |
|  Custom mGenomeWide | Now | Genome regulatory networks | CRISPR i/a  | Highest Quality / Custom Designs |
|  Custom any Organism/Vector | Now | Various, fully custom | Customer design  | Highest Quality / Custom Designs |
|  HR Donor libraries | Now | Donor DNA for knock-ins, fully custom | Genome wide tagging, reporters | Quality/ Design freedom |

Enabling scientist to generate and test genomics hypothesis

Agilent's Advantage

Quality



Guide Representation

- Are all the guides present in the pool?

99.99%

Guide Fidelity

- Are the guides all the correct sequence?

Highest fidelity oligo synthesis

Customization



Custom Content

- What if I have my own set of guides/targets?

Design service and fully custom

Workflow flexibility

- What if I want to delivery guides using my own technique?

Customization at catalog prices

Summary

- **Best performing sureselect Exomes with option to add UTR and additional content**
- **Fastest workflow**
- **Performs well with low amounts of starting DNA for both Fresh frozen and FFPE samples**
- **Easy to customize**
- **Sureselect target enrichment for any species**
- **Halo-HS power of molecular barcodes enables to identify rare mutation with <1% frequency**
- **Sureselect Metyl seq is an ideal tool for genome wide methylation analysis and customizable.**
- **Pooled Sure guide libraries for Human and Mouse**
- **Best quality Custom Oligo Library to fit your need**

Thank You!

