

Statistical Methods for Genome Wide Regional Analysis with Next Generation Sequencing Data

Hao Wu, Emory University

and

Rafael A Irizarry, DFCI/Harvard

Outline

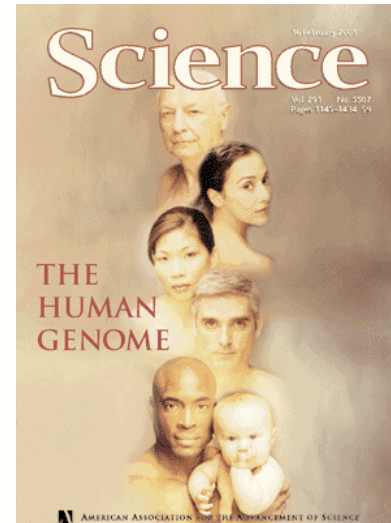
- Introduction to Next Generation Sequencing (NGS)
- Motivation for region finding
- ChIPSeq
- Whole genome bisulfite sequencing (WGBS)
- Computer Lab



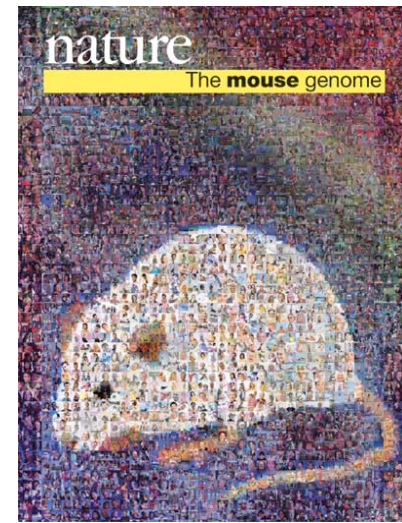
D. melanogaster, Science, 2000



H. sapiens, Nature, 2000



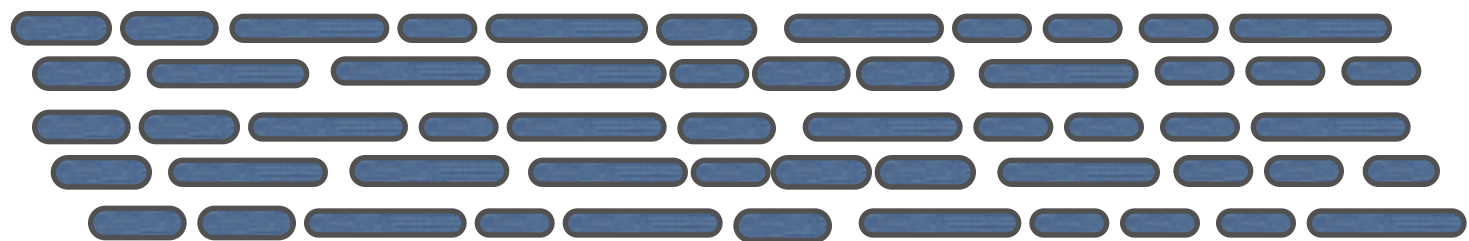
Science, 2000

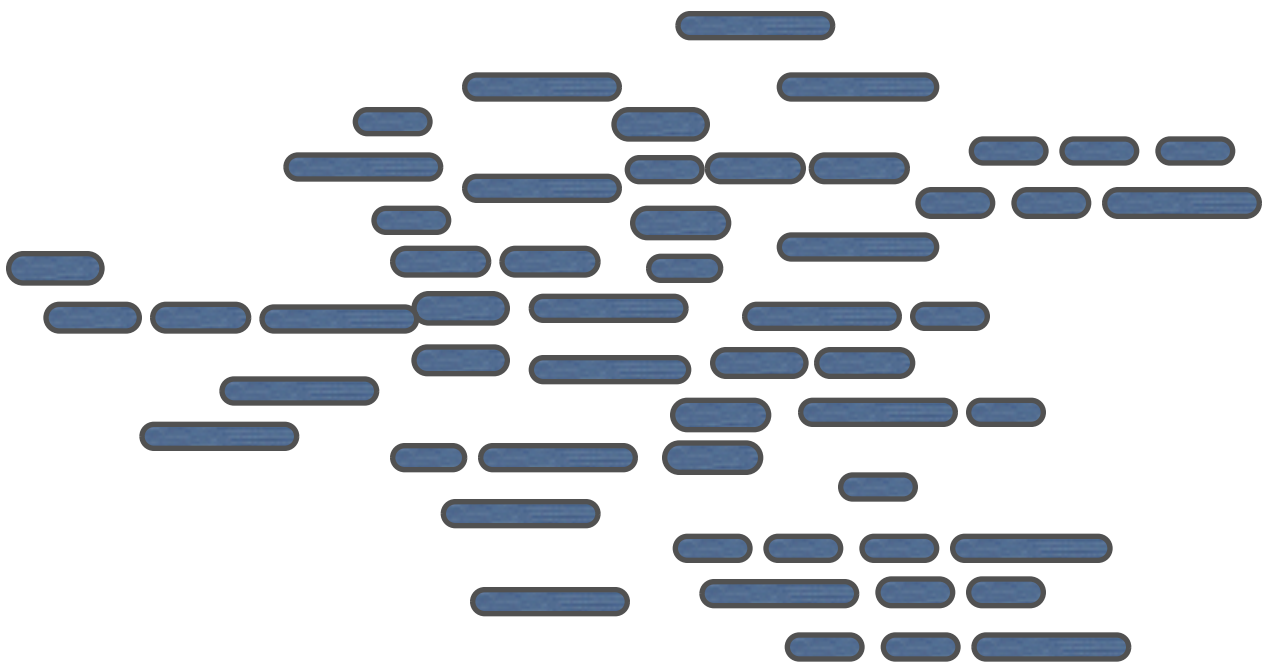


M. musculus, Nature, 2002

- Back then: millions of clones (thousand bps) in 9 months for billions of dollars
- Today: billion of short reads (35-100 bps) in a week for thousands of dollars
- Claim: Assemble a genome in weeks for less than \$100,000







Sequence first 35-400 bps: call them “*reads*”

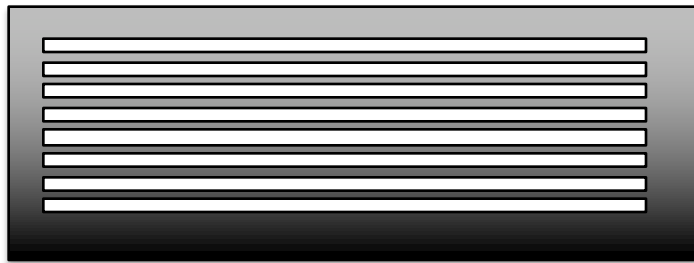
```
GTTGAGGCTTGCGTTTTTGGTACGCTGGACTTTGT  
GTACTCGTCGCTGCGTTGAGGCTTGCGTTTTTGGT  
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC  
CTTGCGTTTATGGTACGCTGGACTTTGTAGGATAC  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC  
GCGTTTATGGTACGCTGGACTTTGTAGGATACCCT  
GAGGCTTGCGTTTATGGTACGCTGGACTTTGTAGG  
GCGTTGAGGCTTGCGTTTATGGTACGCTGGATTTT  
CGTTTATGGTACGCTGGACTTTGTAGGATACCCTC  
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT  
GTTTATGGTACGCTGGACTTTGTAGGATACCCTCG  
TCTCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTA  
TGCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTA  
GCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTAC  
TATGGTACGCTGGACTTTGTAGGATACCCTCGCTT  
TCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTTTG  
CGTCGCTGCGTTGAGGCTTGCGTTTATGGTACGCT  
GTTGAGGCTTGCGTTTATGGTACGCTGGGCTTTTT  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC
```

Available platforms

- Major player:
 - Illumina: HiSeq, MiSeq.
 - LifeTech: SOLiD, IonTorrent.
 - Roche 454.
- Others:
 - Complete Genomics
 - Pacific Bioscience
 - Helicos



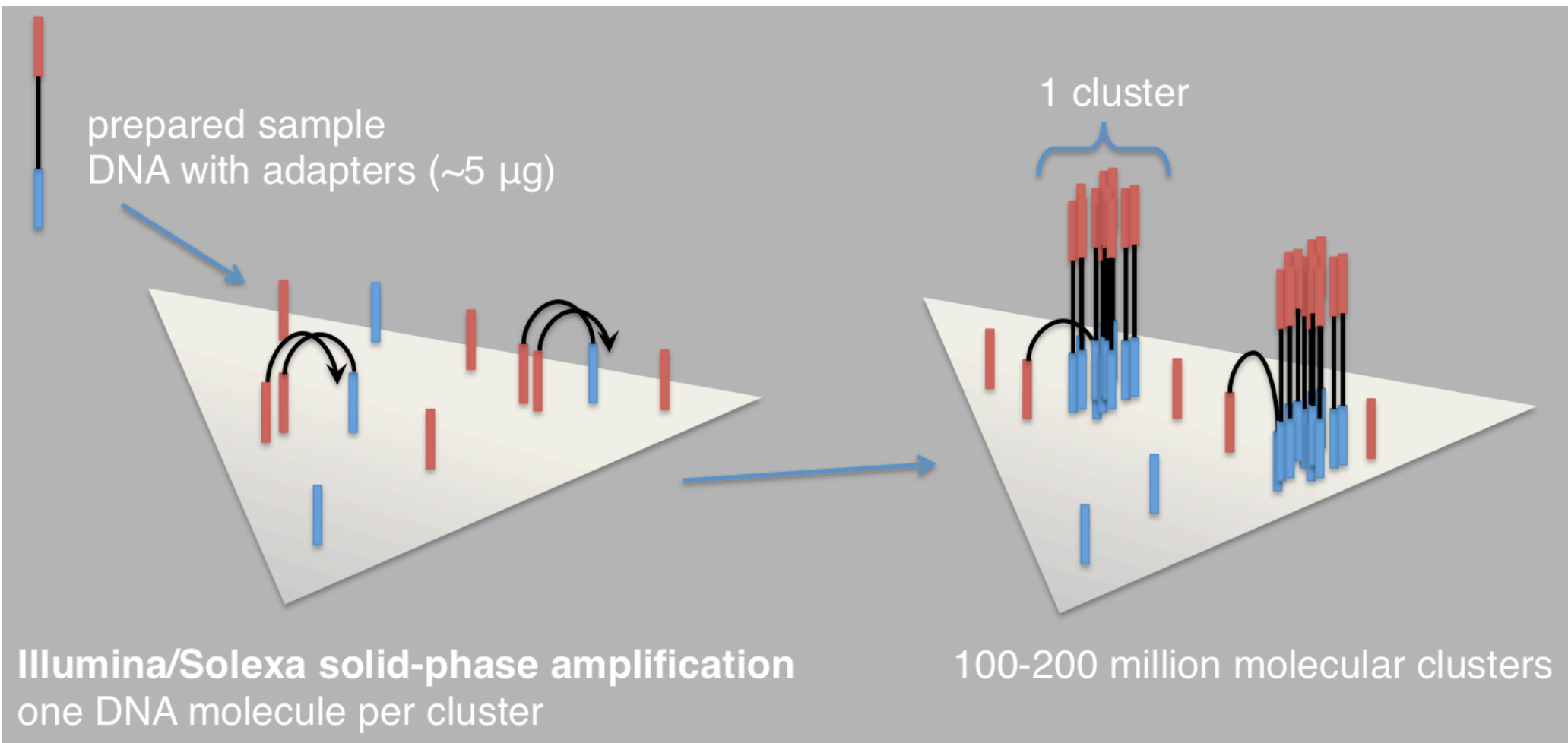
~7 cm



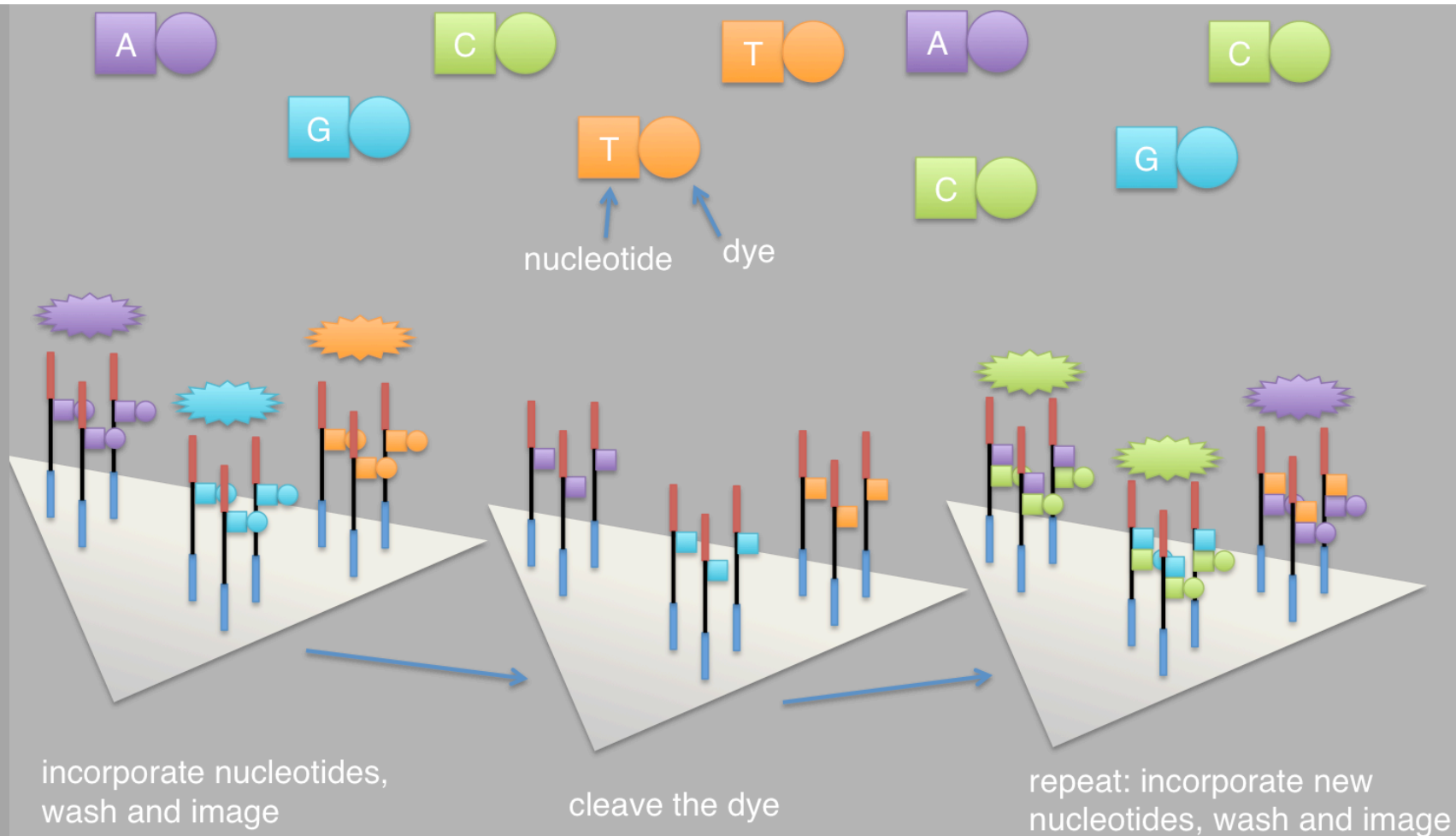
~2 cm

Illumina “flow cell”

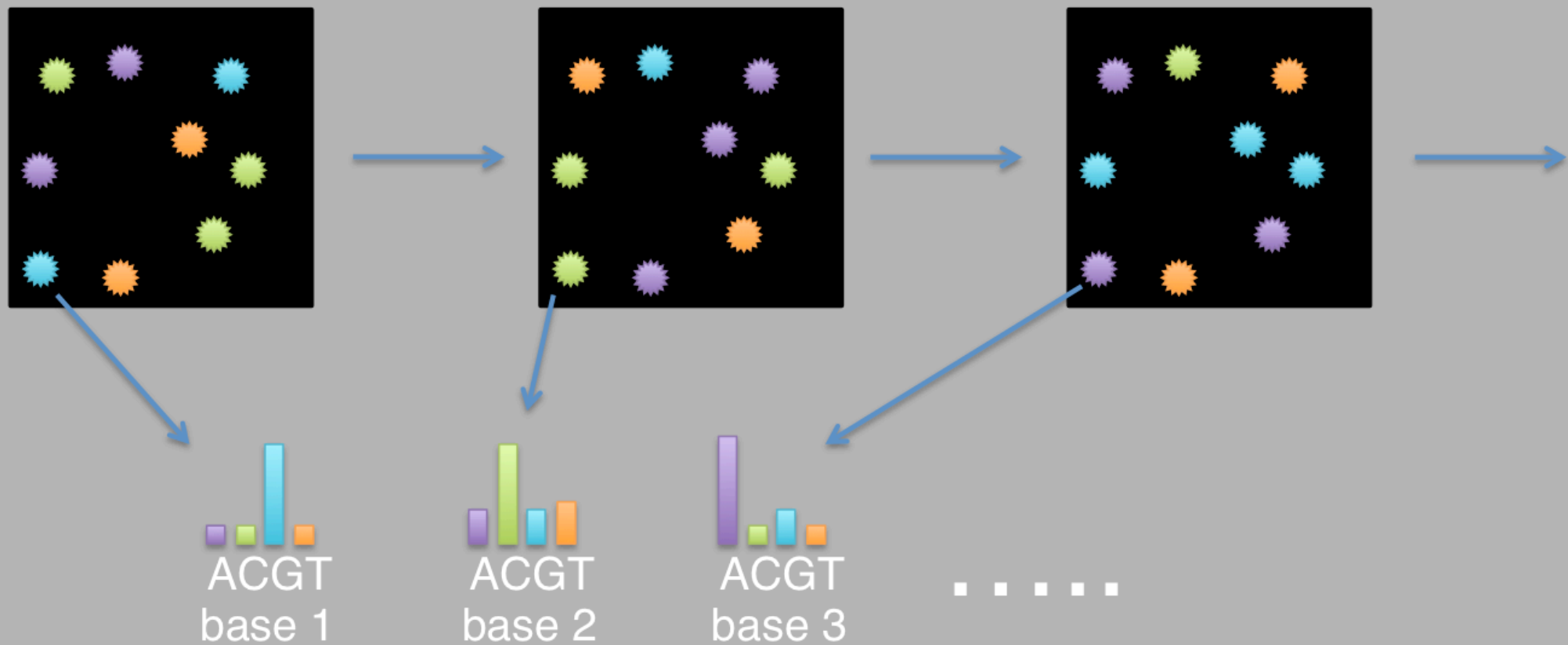
- Eight lanes
- ~160M short reads (~50-70 bp) per lane



Bridge amplification



Sequencing all bases at once



Images from sequencing machine

Raw sequence reads from NGS

- Large text file (millions of lines) with simple format.
 - Most frequently used: fasta/fa format for storing the sequences, or fastq format storing both the sequence and corresponding quality scores.
- fasta format:

read name	→	>5_143_428_832
read sequence	→	GATATTGTAGCATAACGCAACTTGGGAGGTGAGCTT

```
>5_143_984_487
GTTTTCATGCCTCCAAATCTTGGAGGCTTTTTTATG
>5_143_963_690
GGTATATGCACAAAATGAGATGCTTGCTTATCAACA
>5_143_957_461
GGAGGGTGTCAATCCTGACGGTTATTTCTAGACAA
>5_143_808_403
GATAACCGCATCAAGCTCTTGAAGAGATTCTGTCT
```

fastq format

@HWI-EAS165:1:1:50:908:1

CTGCGGTCTCTAAAGTGCCATCTCATTGTGCTTTGTATCAGTCAGTGCTGGA

+

BCCBCB8ABBBBBBB:BC=8@BBA:@BB@BBBCBB<9BBAC;A<C?BAAB<#

@HWI-EAS165:1:1:50:0:1

NCAACCCACAGTAATATGTAAACAAAACTAAACCAGGAGCTGAAGGG

+

#BABABBBBBB@08<@?A@7:A@CCBCCCCBBBCCBB=?BBBB@7@B=A>:2

@HWI-EAS165:1:1:50:708:1

GGTCAGCATGTCTTCTGTAAAGTGCTTGCACAAGCTAGCCTCTGCCTATGGG

+

BB@A;B>@A@@=BB=BB?A>@@>B?ABBA=A?@@>@@A:=?>?A@=B8@@AB

@HWI-EAS165:1:1:50:1494:1

CTGGTGTACACAAGCAGGTCTCCTGTGTTGACTTCACCAGACACTGTCATT

+

BCBB@AB@1ABBBBBBAAB?BBBBAB<A?AA>BB@?1ABBA@BBBA@;B>>:

← read name

← read sequence

← separator

← quality scores

Single-end vs. paired-end sequencing

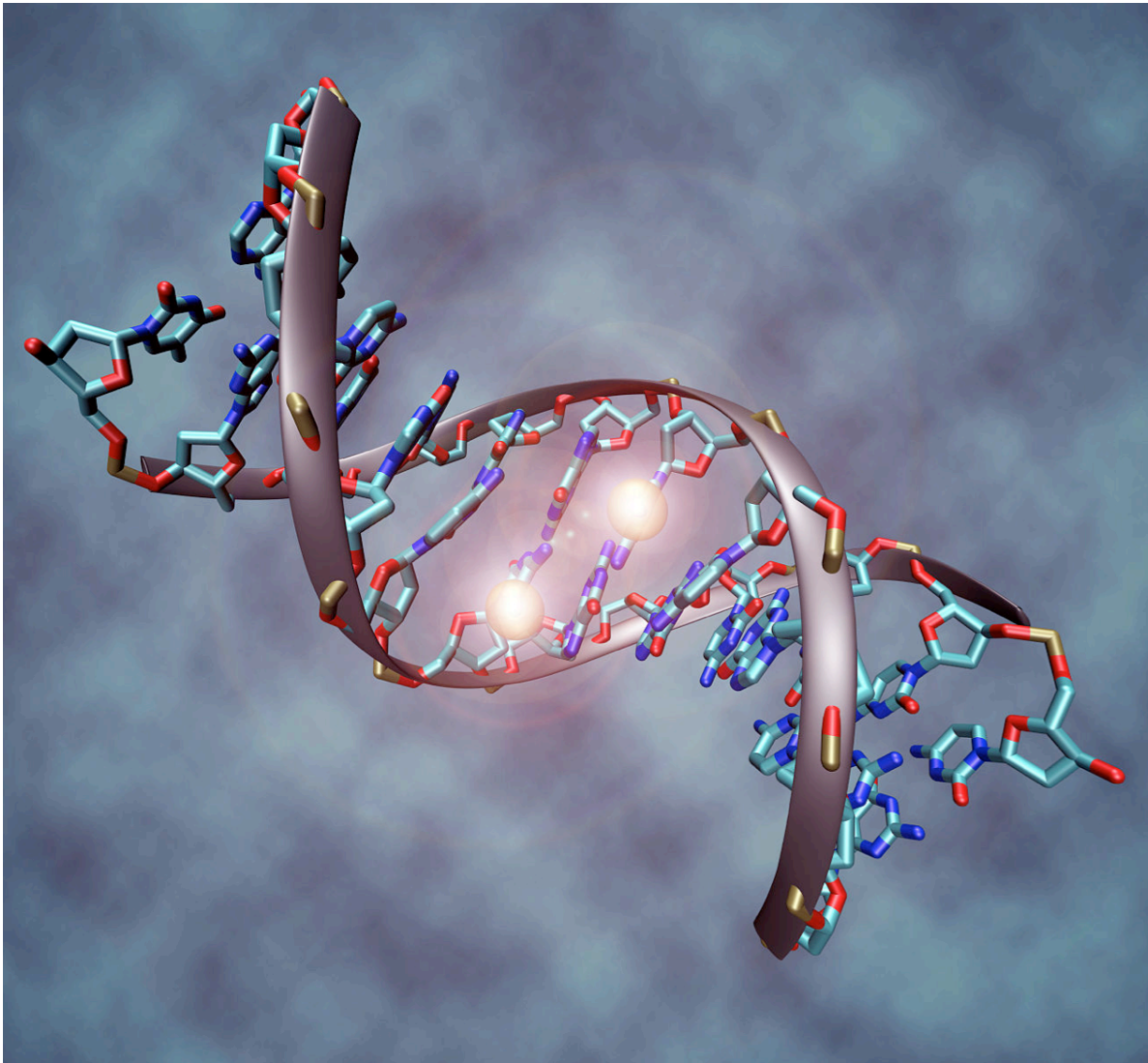
- Sequence one or both ends of the DNA segments.
- **Single-end** sequencing: sequence one end of the DNA segment.
- **Paired-end** sequencing: sequence both ends of a DNA segments.
 - Result reads are “paired”, separated by certain length (the length of the DNA segments, usually a few hundred bps).
 - Paired-end data can be used as single-end, but contain extra information which is useful in some cases, e.g., detecting structural variations in the genome.
 - Modeling technique is more complicated.

Segment 2 – Applications of NGS in Genomics

(do not include this slide in video)

Not just Assembly

- Resequencing
- SNP discovery and genotyping
- Variant discovery and quantification
- TF binding sites: ChIP-Seq
- Gene expression: RNA-Seq
- Measuring methylation



What to do with all these sequences?

```
GTTGAGGCTTGCGTTTTTGGTACGCTGGACTTTGT  
GTACTCGTCGCTGCGTTGAGGCTTGCGTTTTTGGT  
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC  
CTTGCGTTTATGGTACGCTGGACTTTGTAGGATAC  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC  
GCGTTTATGGTACGCTGGACTTTGTAGGATACCCT  
GAGGCTTGCGTTTATGGTACGCTGGACTTTGTAGG  
GCGTTGAGGCTTGCGTTTATGGTACGCTGGATTTT  
CGTTTATGGTACGCTGGACTTTGTAGGATACCCTC  
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT  
GTTTATGGTACGCTGGACTTTGTAGGATACCCTCG  
TCTCGTCGCTCGTCGCTGCGTTGAGGCTTGCGTTTA  
TGCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTA  
GCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTAC  
TATGGTACGCTGGACTTTGTAGGATACCCTCGCTT  
TCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTTTG  
CGTCGCTGCGTTGAGGCTTGCGTTTATGGTACGCT  
GTTGAGGCTTGCGTTTATGGTACGCTGGGCTTTT  
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC
```

Most apps: Start by matching to reference

GTTGAGGCTTGCGTTTTTGGTACGCTGGACTTTGT
GTACTCGTCGCTGCGTTGAGGCTTGCGTTTTTGGT
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC
CTTGCGTTTATGGTACGCTGGACTTTGTAGGATAC
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC
GCGTTTATGGTACGCTGGACTTTGTAGGATACCCT
GAGGCTTGCGTTTATGGTACGCTGGACTTTGTAGG
GCGTTGAGGCTTGCGTTTATGGTACGCTGGATTTT
CGTTTATGGTACGCTGGACTTTGTAGGATACCCTC
ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT
GTTTATGGTACGCTGGACTTTGTAGGATACCCTCG
TCTCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTA
TGCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTA
GCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTAC
TATGGTACGCTGGACTTTGTAGGATACCCTCGCTT
TCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTTTG
CGTCGCTGCGTTGAGGCTTGCGTTTATGGTACGCT
GTTGAGGCTTGCGTTTATGGTACGCTGGGCTTTTT
TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC

CTCTCGTGCTCGTCGCTGCGTTGAGGCTTGCGTTTATGGTACGCTGGACTTTGTAGGATACCCTCGCTTTC

```

@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNNNN
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGACAGGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGATCCTACCACTCGTCCAANNNNNNNNNNNNNNNN
+
A=B7&7:>B@;A>79:<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1087#0/1
CTCTCTCAAGGTCCCGAAGCAGCAGCAANNNNNANTNCTNNNN
+
BBCCCCCBB7CB7C=7>+<=>B7CB7AAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7+<B=7BB5=ABA?B6BBB4BB7BAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAAAGCATGTTGAGTTCGAGCGNNANANCCTGANNNN
+
BBB@77A7>AAB@>7B=7@.>B7B7AAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:1631#0/1
TGGGAACGAGCTACACTCTCCAGGCTCTCCTCCGTNNNNNN
+
BBB@6@BBB@BB@BBB@BAAB@B7;9BB@BA56<B:AAAA
@HWI-EAS146:5:1:2:1420#0/1
CTCAACTCTGACCTTTGGTGATCCACCGCTNGGCCCTCNNNN
+
BBB@:BBB@BABA7:({=AB@>AAA7AB7=AAAAAAAAAAAA
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNN
+
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGACAGGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGATCCTACCACTCGTCCAANNNNNNNNNNNNNNNN
+
A=B7&7:>B@;A>79:<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1087#0/1
CTCTCTCAAGGTCCCGAAGCAGCAGCAANNNNNANTNCTNNNN
+
BBCCCCCBB7CB7C=7>+<=>B7CB7AAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7+<B=7BB5=ABA?B6BBB4BB7BAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAAAGCATGTTGAGTTCGAGCGNNANANCCTGANNNN
+
BBB@77A7>AAB@>7B=7@.>B7B7AAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:2:1631#0/1
TGGGAACGAGCTACACTCTCCAGGCTCTCCTCCGTNNNNNN
+
RRRRRRRRRRRRRRRRRRR7:RRRRR56<R:AAAAAAA

```

TTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCCCTGCCTCATCCTATTATT

Reference

Variant detection

```

@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNNNN
+
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGACAGCAGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGATCTACCACTCGTCCAAANNNNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGTCCCGAGAGCAGCAGCAANNNNNTNNCTNNNN
+
BBCCCCCBB7CB7C=7>=<=>B7CB7AAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7<-B=7BB5=ABA7B6BBB4BB7BAAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCAGGAGAGCAGCATGTTGAGTTCGAGCGNNNANCTGANNNN
+
BBB9@77A7>AAB@7B=7@.>87B7AAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAA
@HWI-EAS146:5:1:2:1631#0/1
TGGGACGACGCTACACTCTCCAGGCTCTNCTCCGTNNNN
+
BBB@6@BBB@BB@BB@BAAB@B7;9BB@B@B<B:AAAA
@HWI-EAS146:5:1:2:1420#0/1
CTCAACTCTGACCTTTGGTGATCCACCGCTTNGGCTTCNNNN
+
BBB@:BBB@B@B@A?:(=AB@>AAA7AB7=AAAAAAAA
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNN
+
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGACAGCAGNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1848#0/1
CTGGACTGATCTACCACTCGTCCAAANNNNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGTCCCGAGAGCAGCAGCAANNNNNTNNCTNNNN
+
BBCCCCCBB7CB7C=7>=<=>B7CB7AAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7<-B=7BB5=ABA7B6BBB4BB7BAAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCAGGAGAGCAGCATGTTGAGTTCGAGCGNNNANCTGANNNN
+
BBB9@77A7>AAB@7B=7@.>87B7AAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAA
@HWI-EAS146:5:1:2:1631#0/1
TGGGACGACGCTACACTCTCCAGGCTCTNCTCCGTNNNN
+
RRR@6@RRR@RRR@RRR@BAABR7:RRR@B@B<B:AAAA

```



```

GTCGCGAGTANCTGTCT
||||||| |||||
GTCGCGAGTATCTGTCT

GGATCTCGGATATACC
||||| |||||
GGATCT-CGATATACC

AATCTGATCTTATTTT
||||| |||||
AATCTGATCTTATTTT

ATATATATATATATAT
||||| |||||
ATATATATATATATAT

TCTCTCCANNAGAGC
||||| |||||
TCTCTCCAGGAGAGC

```

TTTGGTATTTTCGTC TGGGGGGTATGCACGCGATAGCAT TGCAGACGCT GGAGCCGGAGCACCTATGTCGCGAGTATCTGTCTTTGATTCCCTGCCTCATCCTATTATT

Reference

Variant detection

“Pileup” or “Coverage plot”

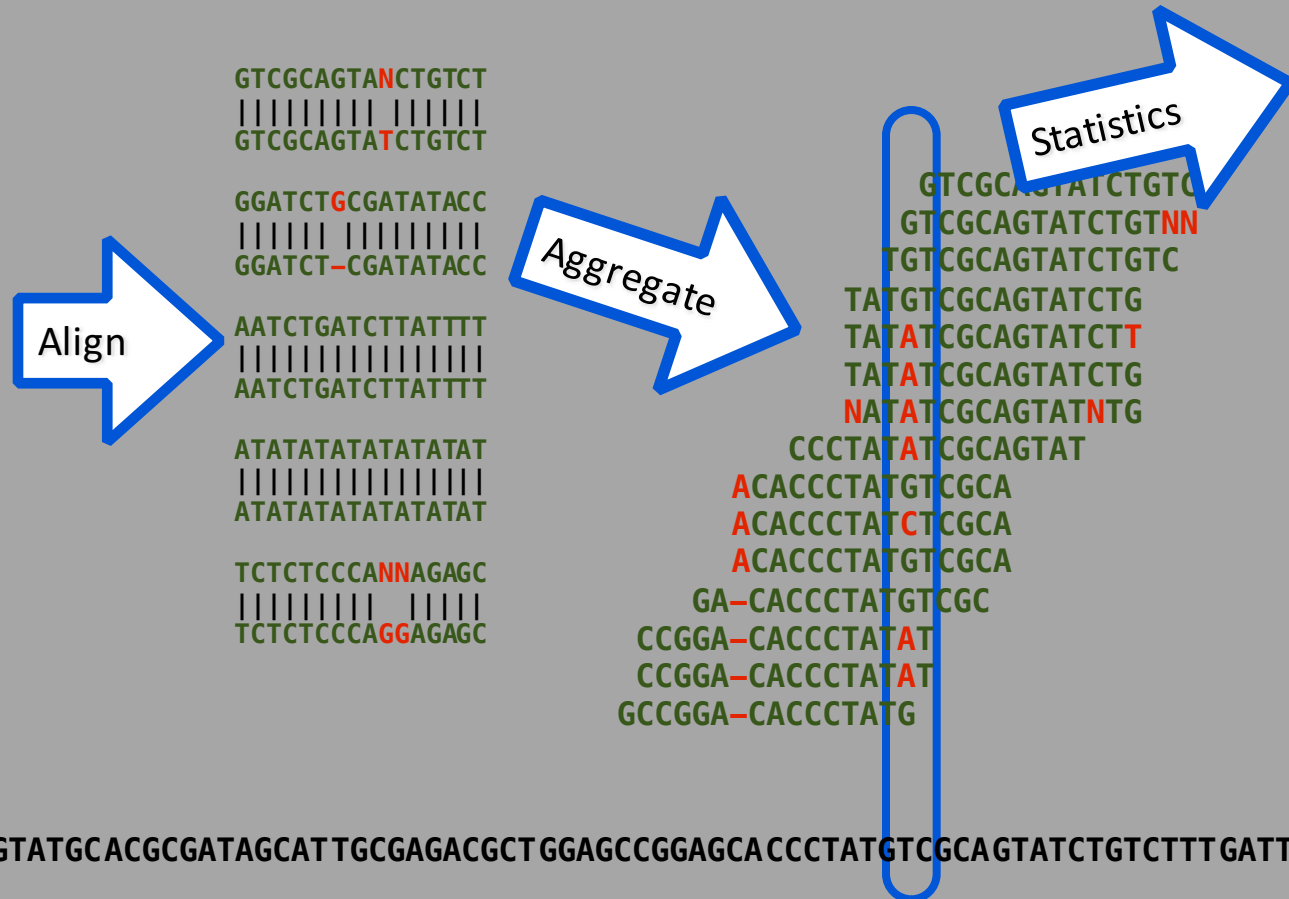
```
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNNNN
+
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAGCTGGTGCAGCAGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGCATCTACCACTCGTCCAAANNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGGTCCCGAGAGCAGCCAAANNNNNTNCTNNNN
+
BBCCCCCBB7CB7C=7>=<=>B7CB7AAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CAGGATCTGGGTTTATTGTAACTCCGCTCANNNGHTNAGNNNN
+
BCC7<-B=7B5=ABA7B6BBB4BB7B>AAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAGCATGTTGAGTTCGAGCCNNANNCCTGANNNN
+
BBB9@77A7>AAB@>7B=7@>87B7AAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAA
TGGGACGAGCCTACACTCTCCAGGCTCTCCTCCGTTNNNN
+
BBB@6@BBB@BB@BB@BAAB@B@7;9BB@BA56<B:AAAA
@HWI-EAS146:5:1:2:1631#0/1
CTCAACTCTGACCTTTGGTGATCCACCCGCTNGGCCCTCNNNN
+
BBB@:BBB@B@A@A@7: (=A@>AAA7AB7=AAAAAAAA
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNN
+
BBBBA7B@;@BBBBA=BA=AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAGCTGGTGCAGCAGNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGCATCTACCACTCGTCCAAANNNNNNNNNNNN
+
A=B767>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGGTCCCGAGAGCAGCCAAANNNNNTNCTNNNN
+
BBCCCCCBB7CB7C=7>=<=>B7CB7AAAAAAAAAAAA
@HWI-EAS146:5:1:1:1719#0/1
CAGGATCTGGGTTTATTGTAACTCCGCTCANNNGHTNAGNNNN
+
BCC7<-B=7B5=ABA7B6BBB4BB7B>AAAAAAAAAAAA
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAGCATGTTGAGTTCGAGCCNNANNCCTGANNNN
+
BBB9@77A7>AAB@>7B=7@>87B7AAAAAAAAAAAA
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>AAAAAAAAAAAA
TGGGACGAGCCTACACTCTCCAGGCTCTCCTCCGTTNNNN
+
RRRRRRRRRRRRRRRRRRR7>RRRRR56<R:AAAA
```



Variant detection

“Pileup” or “Coverage plot”

Call: HET A, G
p-value: 0.0023



Reference

Variant detection

TTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCAGACGCTGGAGCCGGAGCACCTATGTCGAGTATCTGTCTTTGATTCCCTGCCTCATCCTATTATT

“Pileup” or “Coverage plot”

“Depth of coverage” = 14

Call: HET A, G
p-value: 0.0023

```
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA-BA-A~
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGCAGCAGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>~
@HWI-EAS146:5:1:1:1848#0/1
CTGGACTGATCTACCACTCGTCCAAANNNNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747~
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGGTCCCGAGCAGCAGCAANNNNNNNTNNCTNNNN
+
BBCCCCCBB7CB7C=7>~<=BCBCB~
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7~<B=7B5=ABA7B6BBB4BB7B~
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAGCATGTTGAGTTCGAGCCNNANNCCTGANNNN
+
BBB9@77A7>AAB@>7B=7@>87B7~
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>~
TGGGACGAGCCTACACTCTCCAGGCTCTCCTCCGTTNNNN
+
BBB@6@BBB@BB@BBB@BAAB@B@7;9BB@BA56<B~
@HWI-EAS146:5:1:2:1631#0/1
CTCAACTCTGACCTTTGGTGATCCACCCGCTNGGCCCTCNNNN
+
BBB@:BBB@BAA7:~(=AB@>AAA7AB7~A~
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA-BA-A~
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAAGCAGGAAGCTGGTGCAGCAGNNNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>~
@HWI-EAS146:5:1:1:1848#0/1
CTGGACTGATCTACCACTCGTCCAAANNNNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747~
@HWI-EAS146:5:1:1:1687#0/1
CTCTCTCAAGGTCCCGAGCAGCAGCAANNNNNNNTNNCTNNNN
+
BBCCCCCBB7CB7C=7>~<=BCBCB~
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGTNAAGNNNN
+
BCC7~<B=7B5=ABA7B6BBB4BB7B~
@HWI-EAS146:5:1:2:947#0/1
CCGAGGAGAGCATGTTGAGTTCGAGCCNNANNCCTGANNNN
+
BBB9@77A7>AAB@>7B=7@>87B7~
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AAABA77@5AAA:77>~
TGGGACGAGCCTACACTCTCCAGGCTCTCCTCCGTTNNNN
+
RRRRRRRRRRRRRRRRRRRRAARRR7~RRRRRA56<R~
```

Align

```
GTCGCAGTANCTGTCT
||||| |||
GTCGCAGTATCTGTCT

GGATCTCGCATATACC
||||| |||
GGATCT-CGATATACC

AATCTGATCTTATTTT
||||| |||
AATCTGATCTTATTTT

ATATATATATATATAT
||||| |||
ATATATATATATATAT

TCTCTCCANNAGAGC
||||| |||
TCTCTCCAGGAGAGC
```

Aggregate

```
GTCGCAGTATCTGTCT
GTCGCAGTATCTGTNN
TGTCGCAGTATCTGTC
TATGTGCGCAGTATCTG
TATATCGCAGTATCTT
TATATCGCAGTATCTG
NATATCGCAGTATNTG
CCCTATATCGCAGTAT
ACACCCTATGTGCGA
ACACCCTATCTGCGA
ACACCCTATGTGCGA
GA-CACCCTATGTGCG
CCGGA-CACCCTATAT
CCGGA-CACCCTATAT
GCCGGA-CACCCTATG
```

Statistics

“Coverage”

TTTGGTATTTTCGTC TGGGGGGTATGCACGCGATAGCAT TGCAGACGCT GGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCTGCCTCATCCTATTATT

Reference

Variant detection

Sample A

[illegible]

5GAGCTCTCCATGCATTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGAGCCGGAGCACCTATGTGCGAGTATCTGTCTTGATTCTGCCTATCCTATTATTTATCGCACCTA

Gene 1

Sample B

[illegible]

RNA-seq differential expression

Sample A



Align

GTCGCAGTANCTGTCT
||||||| |||||
GTCGCAGTATCTGTCT

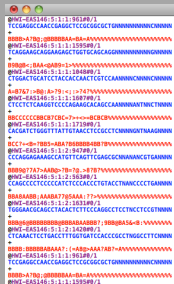
GGATCTGCGATATACC
| | | | | | | | | |
GGATCT-CGATATACC

AATCTGATCTTATTT
|||||||
AATCTGATCTTATTT

5GAGCTCTCCATGCAATTTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCTGCCTATCCTATTATTTATCGCACCTA

Gene 1

Sample B



Align

GTGCGAGTANCTGTCT
||||||| |||||
GTGCGAGTATCTGTCT

GGATCTGCGATATACC
|||||
GGATCT-CGATATACC

AATCTGATCTTATTT
|||||
AATCTGATCTTATTT

RNA-seq differential expression

[illegible]

GTCGCAGTANCTGTCT
 |||||
 GTCGCAGTATCTGTCT
 |||||
 GGATCTGCGATATACC
 |||||
 GGATCT-CGATATACC
 |||||
 AATCTGATCTTATTT
 AATCTGATCTTATTT

TGTCGCAGTATCTGTC
AGCACCCTATGTCGCA
GCCGGAGACCCTATG

[illegible]

GTCGCAGTANCTGTCT
 |||||
 GTCGCAGTATCTGTCT
 |||||
 GGATCTGCGATATACC
 |||||
 GGATCT-CGATATACC
 |||||
 AATCTGATCTTATT
 AATCTGATCTTATT

TGTCGCAGTATCTGTC
AGCACCCTATGTCGCA
GCCGGAGACCCTATG

RNA-seq differential expression

[illegible]

GTCGCAGTANCTGTCT
 |||||
 GTCGCAGTATCTGTCT
 |||||
 GGATCTGCGATATACC
 |||||
 GGATCT-CGATATACC
 |||||
 AATCTGATCTTATTT
 AATCTGATCTTATTT

TGTCGCAGTATCTGTC
AGCACCCTATGTCGCA
GCCGGAGACCCTATG

[illegible]

GTCGCAGTANCTGTCT
 |||||
 GTCGCAGTATCTGTCT
 |||||
 GGATCTGCGATATACC
 |||||
 GGATCT-CGATATACC
 |||||
 AATCTGATCTTATT
 AATCTGATCTTATT

TGTCGCAGTATCTGTC
AGCACCCTATGTCGCA
GCCGGAGACCCTATG

RNA-seq differential expression

```
@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=A\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAAGCAGGAAGAGCTGGTGACAGGNNNNNNNNNNNGNNNNN
+
B9B@B<;BAA<@AB9=1>\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCATCCTACCACCACTCGTCCAANNNNNNNNNNNNNNN
+
A=B7&7:>B@;A>79:<;>747\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1607#0/1
CTCCTCTCAAGGTCCCGACAGCACAGCCAAANNNNNANTHNTNNNN
+
BBCCCCCBB7CB7C@>7>+<=>B7CB\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGNTAAAGNNNN
+
BCT7+<B=7BB5=ABA7B6BBB4BB7B\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:947#0/1
CCGAGGAGAAAGCCATGTTCACTTCGAGCGCNANANCGTGANNNN
+
BBB9@7A7>AAB@>7B=7@.>B7B7\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:563#0/1
CCAGCCCCCTCCCCATCTCCACCTGTACCTNANCCCCGTGANNNN
+
BBABAAB@;AABA77@SAAA:7>\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCCTACACTCTTCCAGGCTCCTNCCTCCGTNNNN
+
BBB@6@BBB@BBB@BBB@AABBB7;9BB@BA5<B:\AAAA
@HWI-EAS146:S:1:2:1420#0/1
CTCAAACTCTGACCTTTGGTGATCCACCCGCTNGGCCCTTCNNNN
+
BBBB;BBBBBABA7:([=A@>AAA7AB7=A\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGCGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=A\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAAGCAGGAAGAGCTGGTGACAGGNNNNNNNNNNNGNNNNN
+
B9B@B<;BAA<@AB9=1>\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCATCCTACCACCACTCGTCCAANNNNNNNNNNNNNNN
+
A=B7&7:>B@;A>79:<;>747\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1607#0/1
CTCCTCTCAAGGTCCCGACAGCACAGCCAAANNNNNANTHNTNNNN
+
BBCCCCCBB7CB7C@>7>+<=>B7CB\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNGNTAAAGNNNN
+
BCT7+<B=7BB5=ABA7B6BBB4BB7B\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:947#0/1
CCGAGGAGAAAGCCATGTTCACTTCGAGCGCNANANCGTGANNNN
+
BBB9@7A7>AAB@>7B=7@.>B7B7\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:563#0/1
CCAGCCCCCTCCCCATCTCCACCTGTACCTNANCCCCGTGANNNN
+
BBABAAB@;AABA77@SAAA:7>\AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCCTACACTCTTCCAGGCTCCTNCCTCCGTNNNN
+
BBB@6@BBB@BBB@BBB@AABBB7;9BB@BA5<B:\AAAA
```

TTGGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCCCTGCCTCGGGAGCTCTCCA

Reference

ChIP-seq

```

@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAAGCAGGAAGAGCTGGTGACAGGNNNNNNNNNNNNNNNN
+
BBD@B<;BAA<@B@=1>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCATCTACCACTCGTCAANNNNNNNNNNNNNNNNNN
+
A=B7&7>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1607#0/1
CTCTCTCAAGGTCCTCCAGAGCACAGCAANNNNNNNNNNNNNN
+
BBCCCCCBBB7C@=7><=>B7CBB7AAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNNNNNNNNNN
+
BCT7<-B=7BB5=ABA7B6BBB4BB7BAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:947#0/1
CCAGGAGAAAGCCTGTTCACTTCGAGCGCANNANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7AAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCCGANNNN
+
BBABAAB@;AABA77@SAAA:7>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCCTACACTCTCCAGGCTCCTTNCCTCCGTNNNN
+
BBB@G@BBB@BBB@BBB@AABBB7;9BB@B45<B:AAAA
@HWI-EAS146:S:1:2:1420#0/1
CTCAAACTCTGACCTTTGGTGATCCACCGCTTNGGCTTNNNN
+
BBB@;BBB@BABA7:([=A@>AAA7AB7=AAAAAAAAAAAA
@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAAGCAGGAAGAGCTGGTGACAGGNNNNNNNNNNNNNNNN
+
BBD@B<;BAA<@B@=1>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCATCTACCACTCGTCAANNNNNNNNNNNNNNNNNN
+
A=B7&7>B@;A>79<;>747AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1607#0/1
CTCTCTCAAGGTCCTCCAGAGCACAGCAANNNNNNNNNNNNNN
+
BBCCCCCBBB7C@=7><=>B7CBB7AAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTTATTGTAACTCCGCTCNNNNNNNNNNNNNN
+
BCT7<-B=7BB5=ABA7B6BBB4BB7BAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:947#0/1
CCAGGAGAAAGCCTGTTCACTTCGAGCGCANNANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7AAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCCGANNNN
+
BBABAAB@;AABA77@SAAA:7>AAAAAAAAAAAAAAAAAAAA
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCCTACACTCTCCAGGCTCCTTNCCTCCGTNNNN
+
RRR@G@RRR@RRR@RRR@AABBB7-@RR@B45<B:AAAA

```



```

GTCGCAGTANCTGTCT
||||||| |||||
GTCGCAGTATCTGTCT

GGATCTGCGATATACC
||||| |||||
GGATCT-CGATATACC

AATCTGATCTTATTTT
||||||| |||||
AATCTGATCTTATTTT

ATATATATATATATAT
||||||| |||||
ATATATATATATATAT

TCTCTCCCANNAGAGC
||||||| |||||
TCTCTCCCAAGAGAGC

```

TTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCCCTGCCTCGGGAGCTCTCCA

Reference

ChIP-seq


```

@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7BQ;@BBBBBAA=BA=A~~~~~
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAGCAGGAAGAGCTGGTGACAGCAGNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>~~~~~
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGCTCTACCACTCGTCCAAANNNNNNNNNNNNNN
+
A=B7&7>BQ;A>79<;>747~~~~~
@HWI-EAS146:5:1:1:1607#0/1
CTCTCTCAAGGTCACGAGCAGCAANNNNNANTHNCNNNN
+
BBCCCCCBB7C=7><=B7C~~~~~
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTATTGTAACTCCGCTCNNNNGNTAAGNNNN
+
BCT7<-B=7B5=ABA7B6BBB4BB7B~~~~~
@HWI-EAS146:5:1:2:947#0/1
CCAGGAGAAAGCCTGTTCACTTCGAGCGCNNANANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7~~~~~
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCCGANNNN
+
BBABAAB;AABA77@SAA:7>~~~~~
@HWI-EAS146:5:1:2:1631#0/1
TGGGAACGAGCCTACACTCTCCAGGCTCCTNCCTCCGNNNN
+
BBB@G@BBB@BBB@BBB@AABBB7;9BB@B45<B:~~~~~
@HWI-EAS146:5:1:2:1420#0/1
CTCAACTCTGACCTTTGGTGATCCACCGCTTNGGCTTCNNNN
+
BBB;BBB@BABA7:([A@>AAA7AB7=A~~~~~
@HWI-EAS146:5:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNN
+
BBBB-A7BQ;@BBBBBAA=BA=A~~~~~
@HWI-EAS146:5:1:1:1595#0/1
TCAGGAGCAGGAAGAGCTGGTGACAGCAGNNNNNNNNNNNNNN
+
B9B@B<;BAA<@B9=1>~~~~~
@HWI-EAS146:5:1:1:1048#0/1
CTGGACTGCTCTACCACTCGTCCAAANNNNNNNNNNNNNN
+
A=B7&7>BQ;A>79<;>747~~~~~
@HWI-EAS146:5:1:1:1607#0/1
CTCTCTCAAGGTCACGAGCAGCAANNNNNANTHNCNNNN
+
BBCCCCCBB7C=7><=B7C~~~~~
@HWI-EAS146:5:1:1:1719#0/1
CACGATCTGGGTTATTGTAACTCCGCTCNNNNGNTAAGNNNN
+
BCT7<-B=7B5=ABA7B6BBB4BB7B~~~~~
@HWI-EAS146:5:1:2:947#0/1
CCAGGAGAAAGCCTGTTCACTTCGAGCGCNNANANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7~~~~~
@HWI-EAS146:5:1:2:563#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCCGANNNN
+
BBABAAB;AABA77@SAA:7>~~~~~
@HWI-EAS146:5:1:2:1631#0/1
TGGGAACGAGCCTACACTCTCCAGGCTCCTNCCTCCGNNNN
+
RRR@G@RRR@RRR@RRR@AABBB7-@RR@B45<B:~~~~~

```

Align

```

GTCGCGAGTANCTGTCT
|||||
GTCGCGAGTATCTGTCT

GGATCTGCGATATAACC
|||||
GGATCT-CGATATAACC

AATCTGATCTTATTTT
|||||
AATCTGATCTTATTTT

ATATATATATATATAT
|||||
ATATATATATATATAT

TCTCTCCCANNAGAGC
|||||
TCTCTCCCAAGGAGAGC

```

Aggregate

```

GTCGCGAGTATCTGTCT
GTCGCGAGTATCTGTCT
TGTCGCGAGTATCTGTC
TATGTCGCGAGTATCTG
TATATCGCAGTATCTG
TATATCGCAGTATCTG
TATATCGCAGTATCTG
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
AGCACCCCTATGTCGCA
AGCACCCCTATATCGCA
AGCACCCCTATGTCGCA
GAGCACCCCTATGTCGC
CCGAGCACCCCTATAT
CCGGAGCACCCCTATAT
GCCGGAGCACCCCTATG

```

GATAGCATTGCGAGAC
TATGCACGCGATAGCA

GATTCCTGCCTC

TTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCTATGTCGCGAGTATCTGTCTTTGATTCCTGCCTCGGGAGCTCTCCA

Reference

ChIP-seq

```

@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=AVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAGCAGGAAGAGCTGGTGACAGCAGNNNNNNNNNNNNNN
+
B9B@B<;BAA<AB9=1>VVVVVVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCTCTACCACTCGTCCAAANNNNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1607#0/1
CTCTCTCAAGGTCAGGACGACCAANNNNNNNNNNNNNNNN
+
BBCCCCCBB7CB=7><=>B7B7VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTATTGTAACTCCGCTCNNNNNNNNNNNNNN
+
BCT7<-B=7B5=ABA7B6BBB4BB7B>VVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:947#0/1
CCAGGAGAAAGCATGTTCACTTCGAGCGCANNANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7VVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:1631#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AABA77@SAA:7>VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCTACACTCTCCAGGCTCCTNCTCCGNNNN
+
BBB@G@BBB@BBB@BBB@AAB@B7;9BB@B45<B:VVVVVV
@HWI-EAS146:S:1:2:1420#0/1
CTCAACTCTGACCTTTGGTGATCCACCCGCTNGGCTTNNNN
+
BBB;BBB@B@A@A7:([>A@>AAA7AB7=AVVVVVVVVVVV
@HWI-EAS146:S:1:1:961#0/1
TCCGAGGCCAACCGAGGCTCCGGCGCTGNNNNNNNNNNNNNN
+
BBBB-A7B@;@BBBBBAA=BA=AVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1595#0/1
TCAGGAGCAGGAAGAGCTGGTGACAGCAGNNNNNNNNNNNN
+
B9B@B<;BAA<AB9=1>VVVVVVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1048#0/1
CTGGACTGCTCTACCACTCGTCCAAANNNNNNNNNNNNNN
+
A=B767>B@;A>79<;>747VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1607#0/1
CTCTCTCAAGGTCAGGACGACCAANNNNNNNNNNNNNNNN
+
BBCCCCCBB7CB=7><=>B7B7VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:1:1719#0/1
CACGATCTGGGTTATTGTAACTCCGCTCNNNNNNNNNNNNNN
+
BCT7<-B=7B5=ABA7B6BBB4BB7B>VVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:947#0/1
CCAGGAGAAAGCATGTTCACTTCGAGCGCANNANCTGANNNN
+
BBB9@7A7>AAB@=7B=7@.>B7B7VVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:1631#0/1
CCAGCCCCCTCCCATCTCCACCTGTACCTNANCCCTGANNNN
+
BBABAAB;AABA77@SAA:7>VVVVVVVVVVVVVVVVVVVV
@HWI-EAS146:S:1:2:1631#0/1
TGGGAACGAGCTACACTCTCCAGGCTCCTNCTCCGNNNN
+
RRR@G@RRR@RRR@RRR@A@B@B7;9BB@B45<B:VVVVVV

```

Align

```

GTCGCAGTANCTGTCT
|||||||
GTCGCAGTATCTGTCT

GGATCTGCGATATACC
|||||||
GGATCT-CGATATACC

AATCTGATCTTATTTT
|||||||
AATCTGATCTTATTTT

ATATATATATATATAT
|||||||
ATATATATATATATAT

TCTCTCCCANNAGAGC
|||||||
TCTCTCCCAAGAGAGC

```

Aggregate

```

GTCGCAGTATCTGTCT
GTCGCAGTATCTGTCT
TGTCGCAGTATCTGTC
TATGTCGCAGTATCTG
TATATCGCAGTATCTG
TATATCGCAGTATCTG
TATATCGCAGTATCTG

CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
CCCTATATCGCAGTAT
AGCACCCTATGTCGCA
AGCACCCTATATCGCA
AGCACCCTATGTCGCA
GAGCACCCTATGTCGC
CCGAGCACCCTATAT
CCGGAGCACCCTATAT
GCGGAGCACCCTATG

```

Statistics

Binding occurs here
p-value:
0.0023

```

GATAGCATTGCGAGAC
TATGCACGCGATAGCA
TTGGTATTTTCGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCTATGTCGCAGTATCTGTCTTTGATTCTGCCTCGGGAGCTCTCCA

```

Reference

ChIP-seq