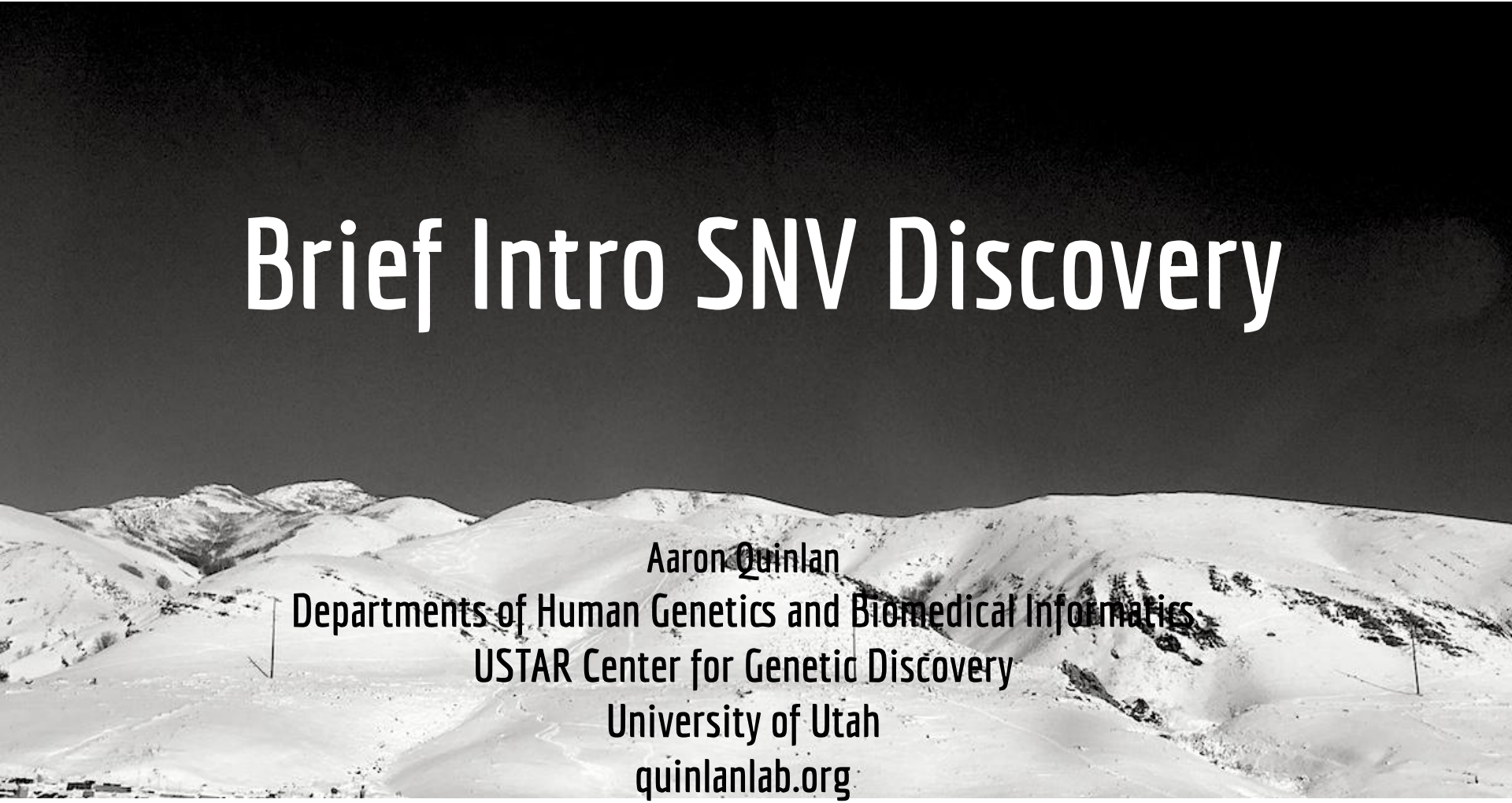
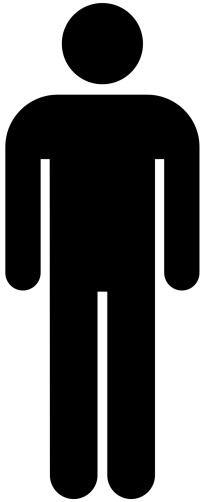
# Brief Intro SNV Discovery

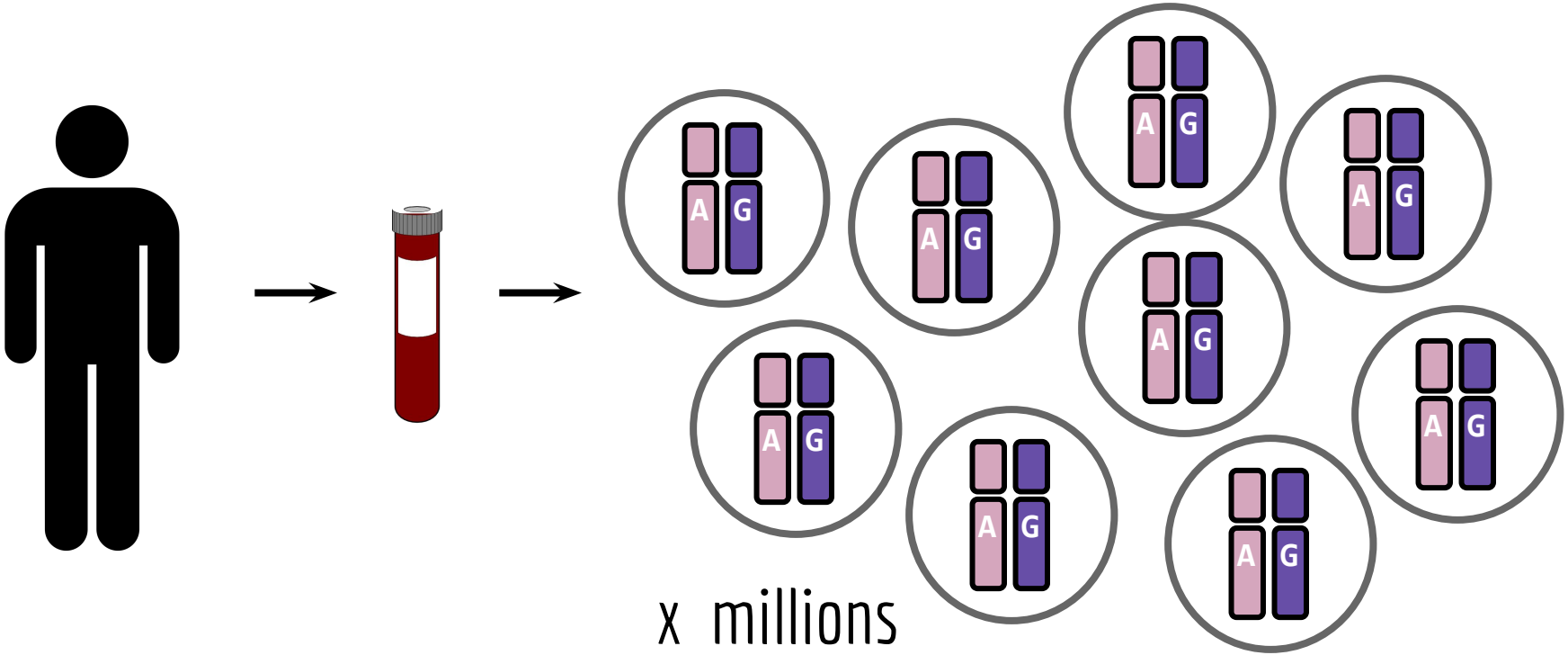


Aaron Quinlan
Departments of Human Genetics and Biomedical Informatics
USTAR Center for Genetic Discovery
University of Utah
quinlanlab.org

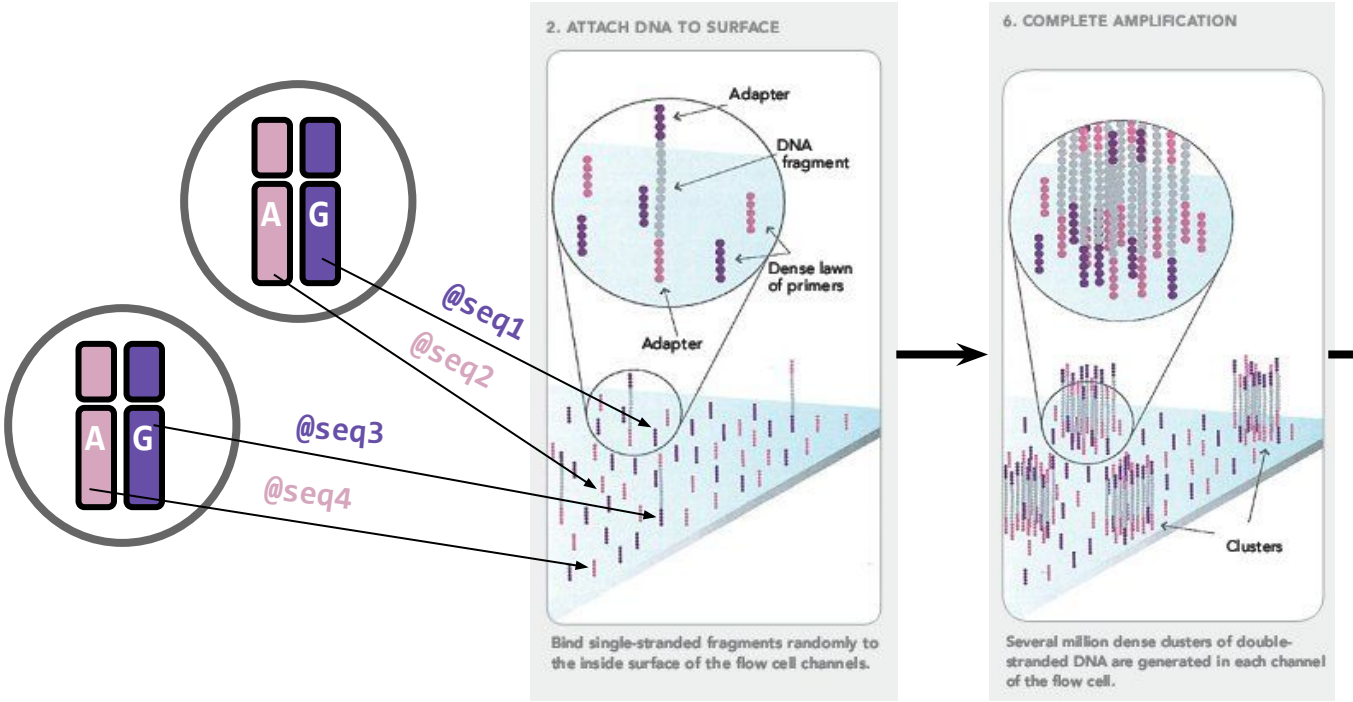
Goal: find all inherited variants in an individual's diploid genome.



# Find inherited genetic variation by sequencing DNA from millions of cells

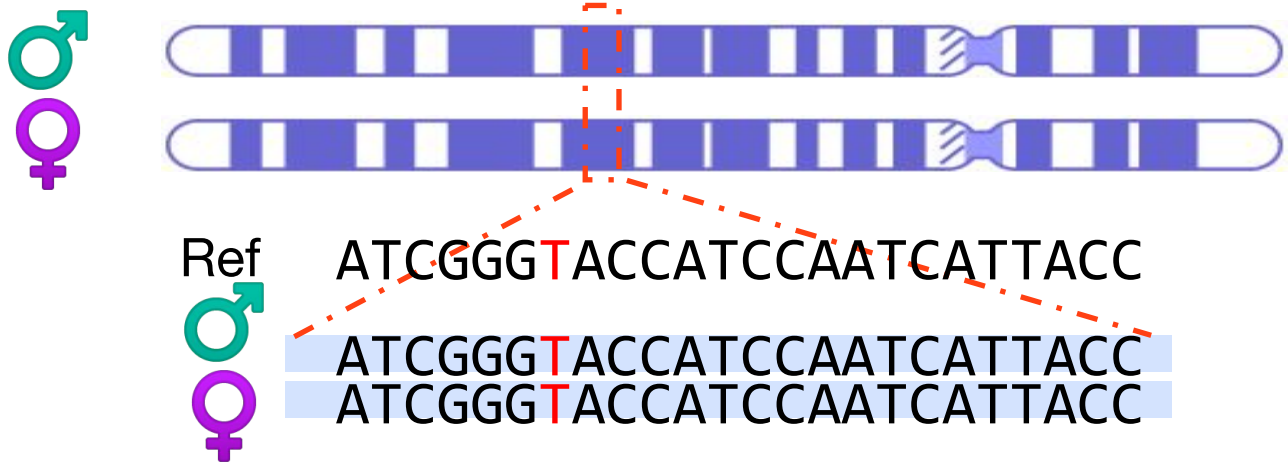


# Each DNA cluster is amplified from a single strand from a single haploid chromosome from a single cell.

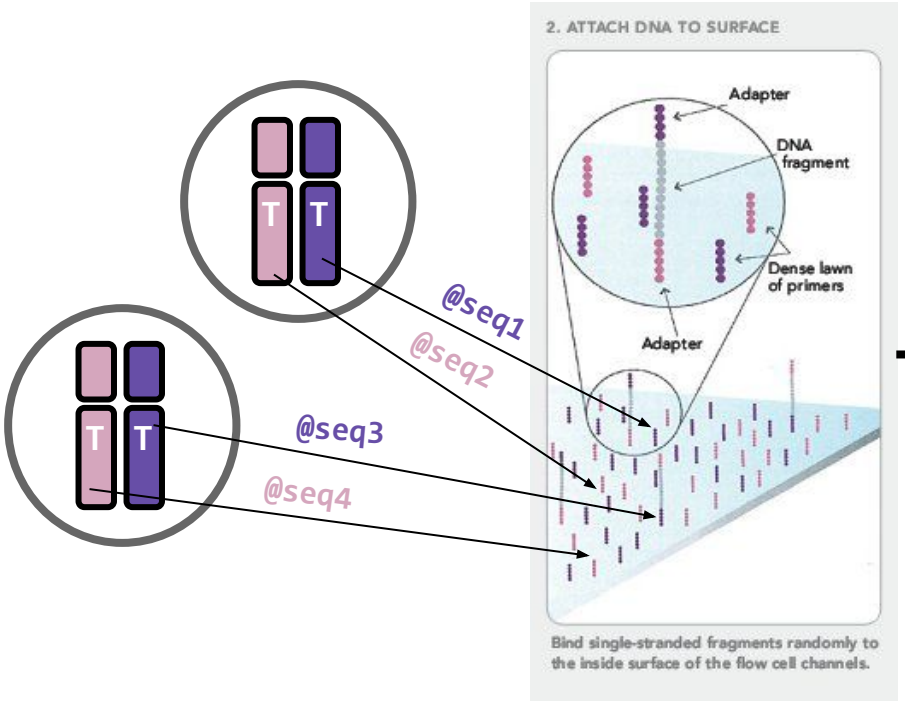



```
@seq1
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq2
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
@seq3
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq4
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
```

# Scenario 1: An individual is homozygous for the "reference" allele.

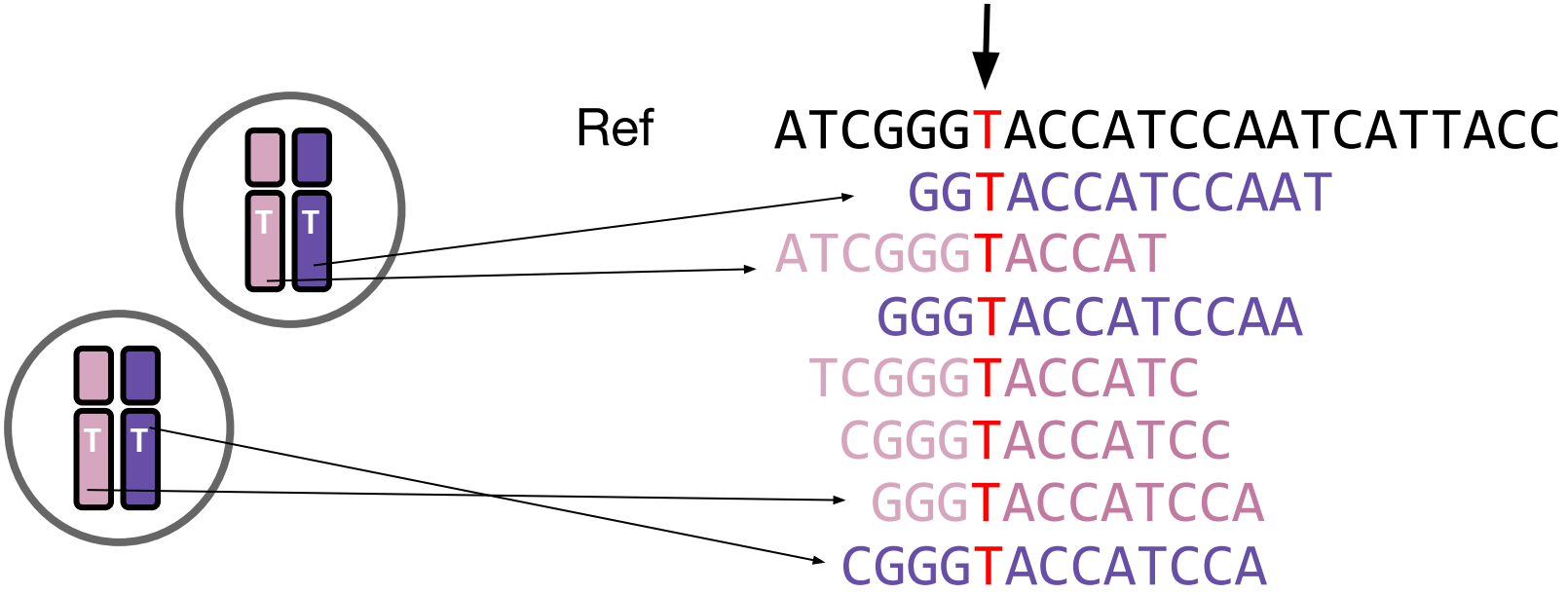


# Scenario 1: An individual is homozygous for the "reference" allele.

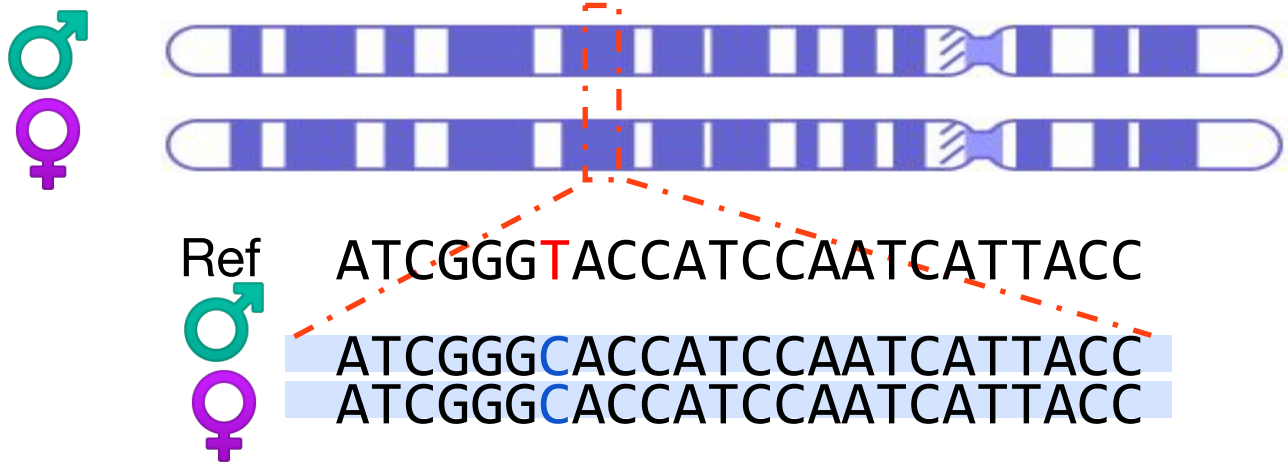



```
@seq1
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq2
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
@seq3
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq4
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
```

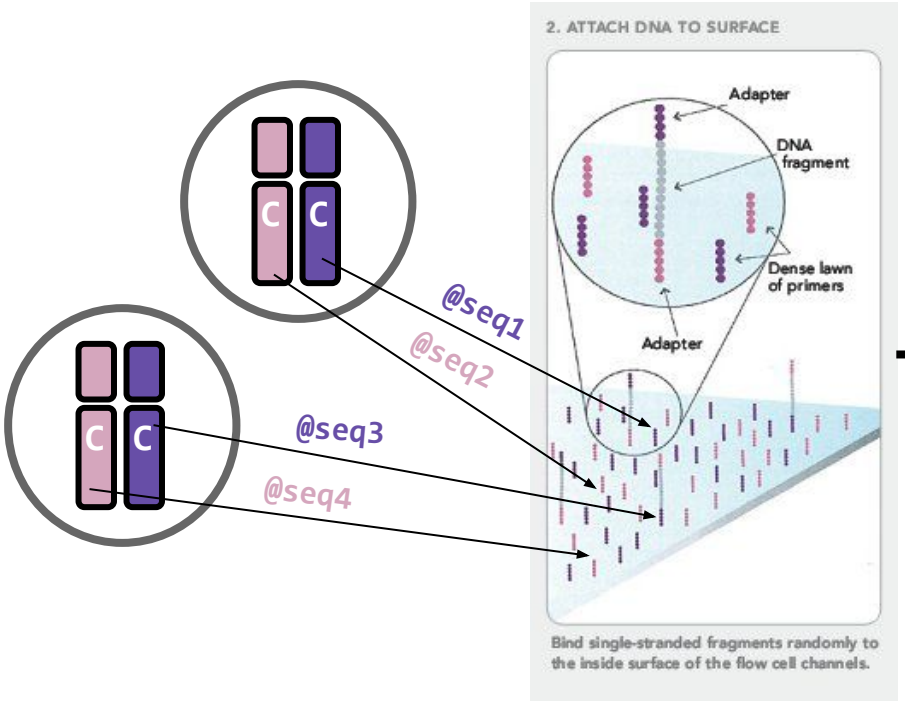
# Scenario 1: An individual is homozygous for the "reference" allele.



# Scenario 2: An individual is homozygous for an "alternate" allele.

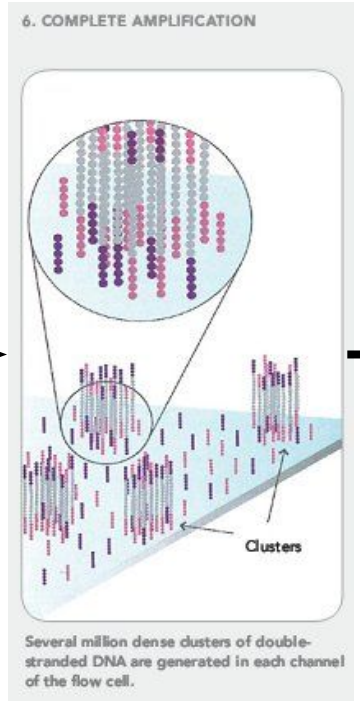


# Scenario 2: An individual is homozygous for an "alternate" allele.



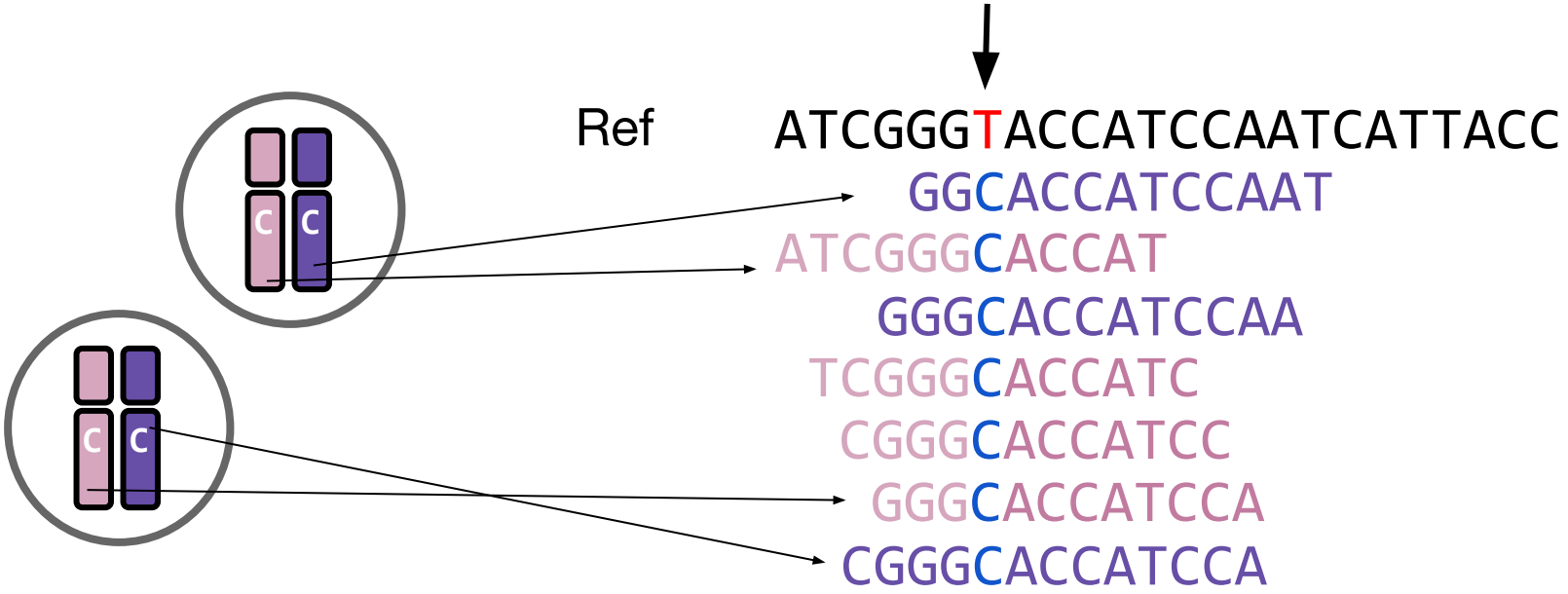




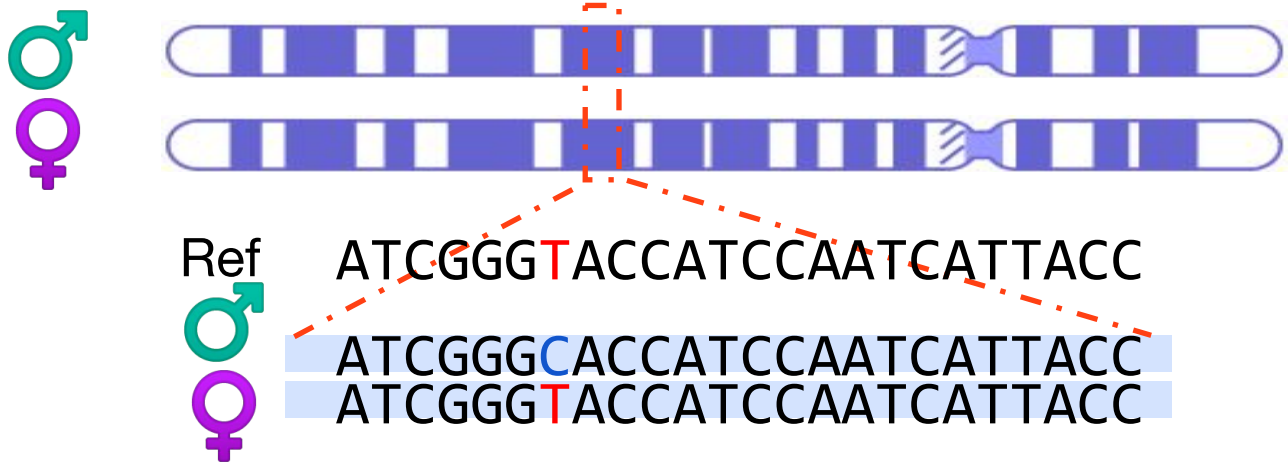



```
@seq1
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq2
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
@seq3
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq4
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
```

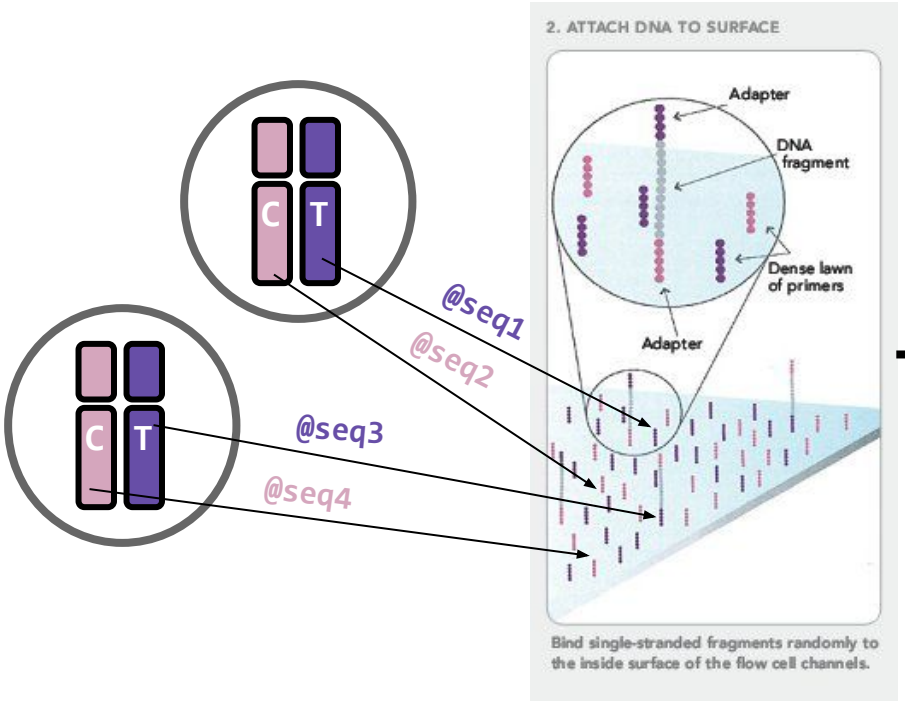
# Scenario 2: An individual is homozygous for an "alternate" allele.



# Scenario 3: An individual is heterozygous for an "alternate" allele.

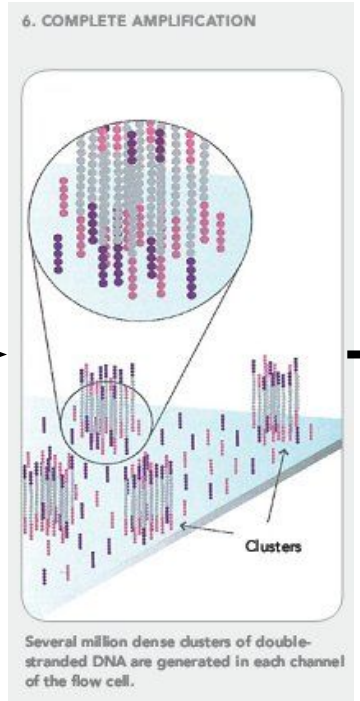


# Scenario 3: An individual is heterozygous for an "alternate" allele.



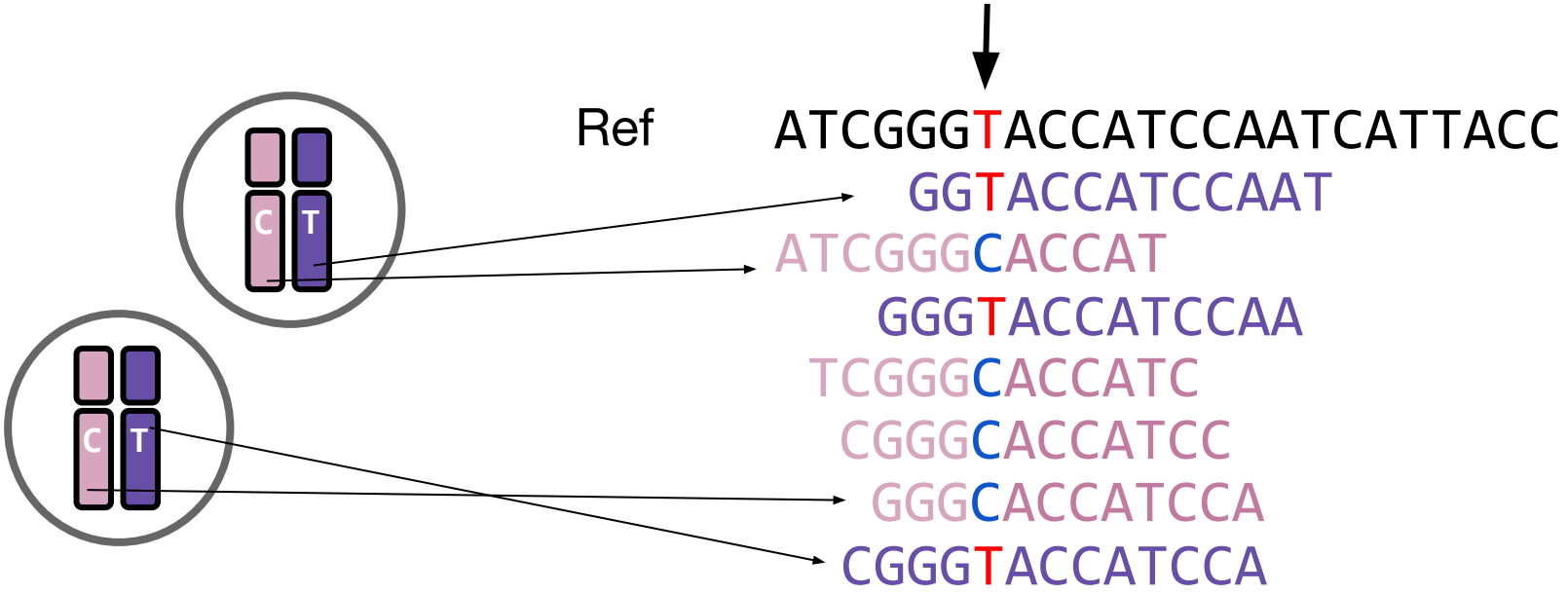







```
@seq1
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq2
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
@seq3
ACCTTCGAACGGCGGGGGGTTACAA
+
!''*((((***+))%%%++).1***
@seq4
TGGAACCGAACGGCCCCGGTTACAT
+
!''*!!!!***+))+++++).1***
```

# Scenario 3: An individual is heterozygous for an "alternate" allele.



# Why might finding heterozygous variants be harder?

# The binomial distribution: adventures in coin flipping



P(heads) = 0.5

P(tails) = 0.5

# Scenario 3: An individual is heterozygous for an "alternate" allele.