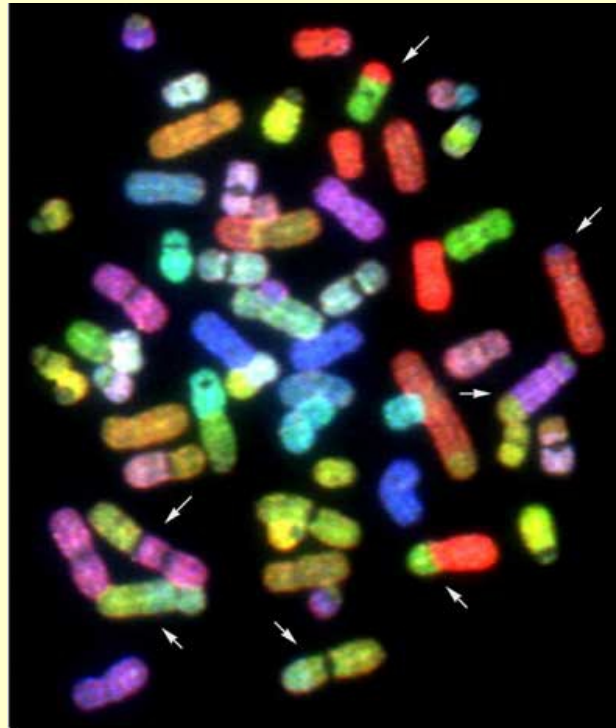


Next Generation Sequencing and Human Genome Databases

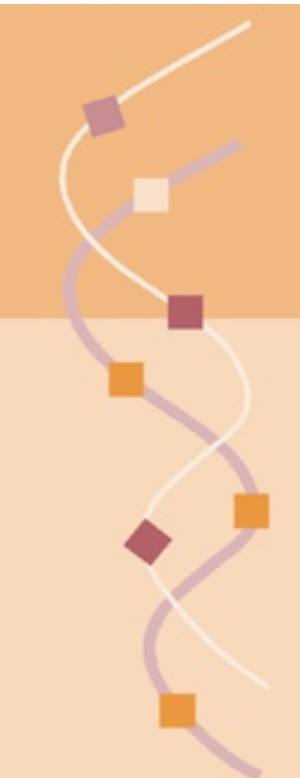
Genomics, Bioinformatics & Medicine
<http://biochem158.stanford.edu/>



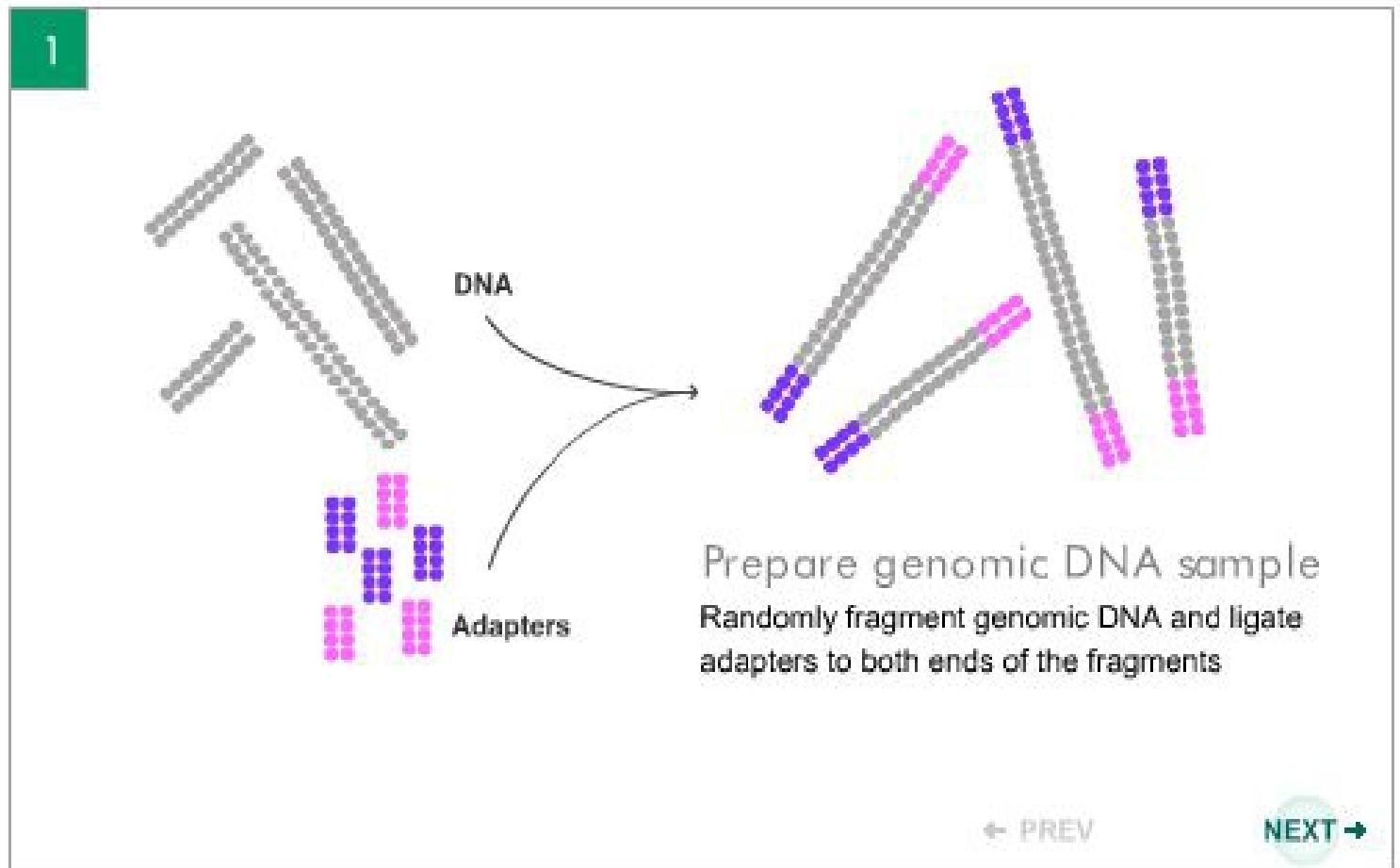
Doug Brutlag

Professor Emeritus of Biochemistry & Medicine
Stanford University School of Medicine

Illumina Solexa Sequencing Technology



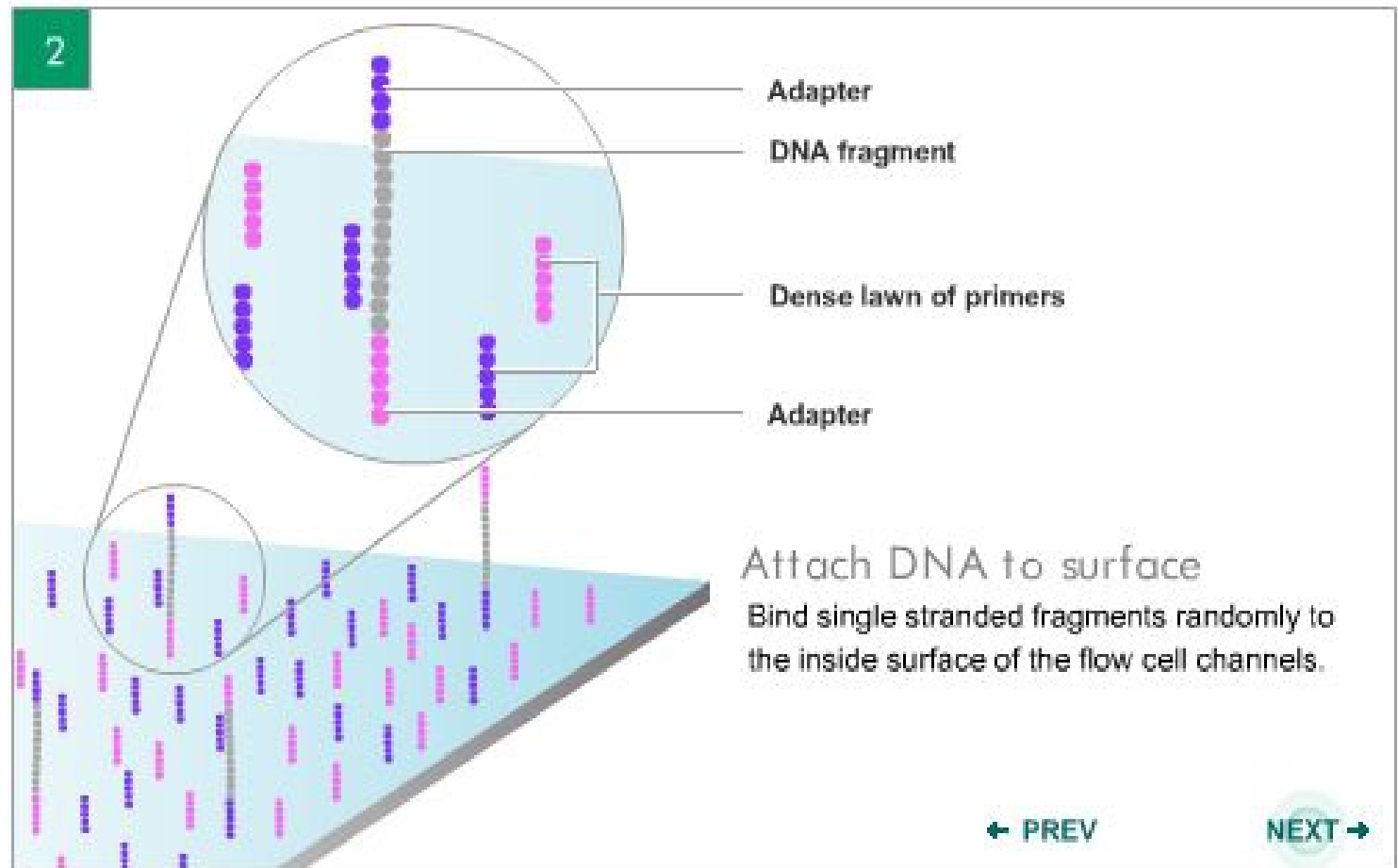
Sequencing-By-Synthesis Demo



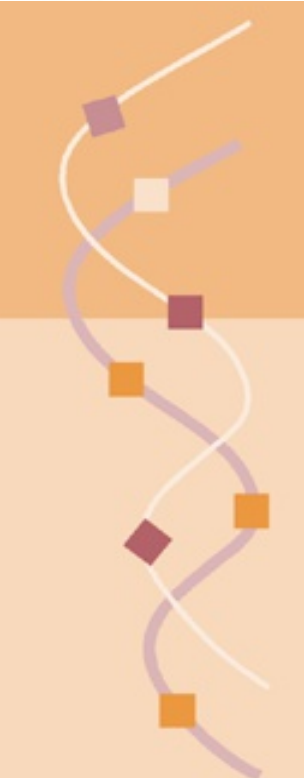
Illumina Solexa Sequencing Technology



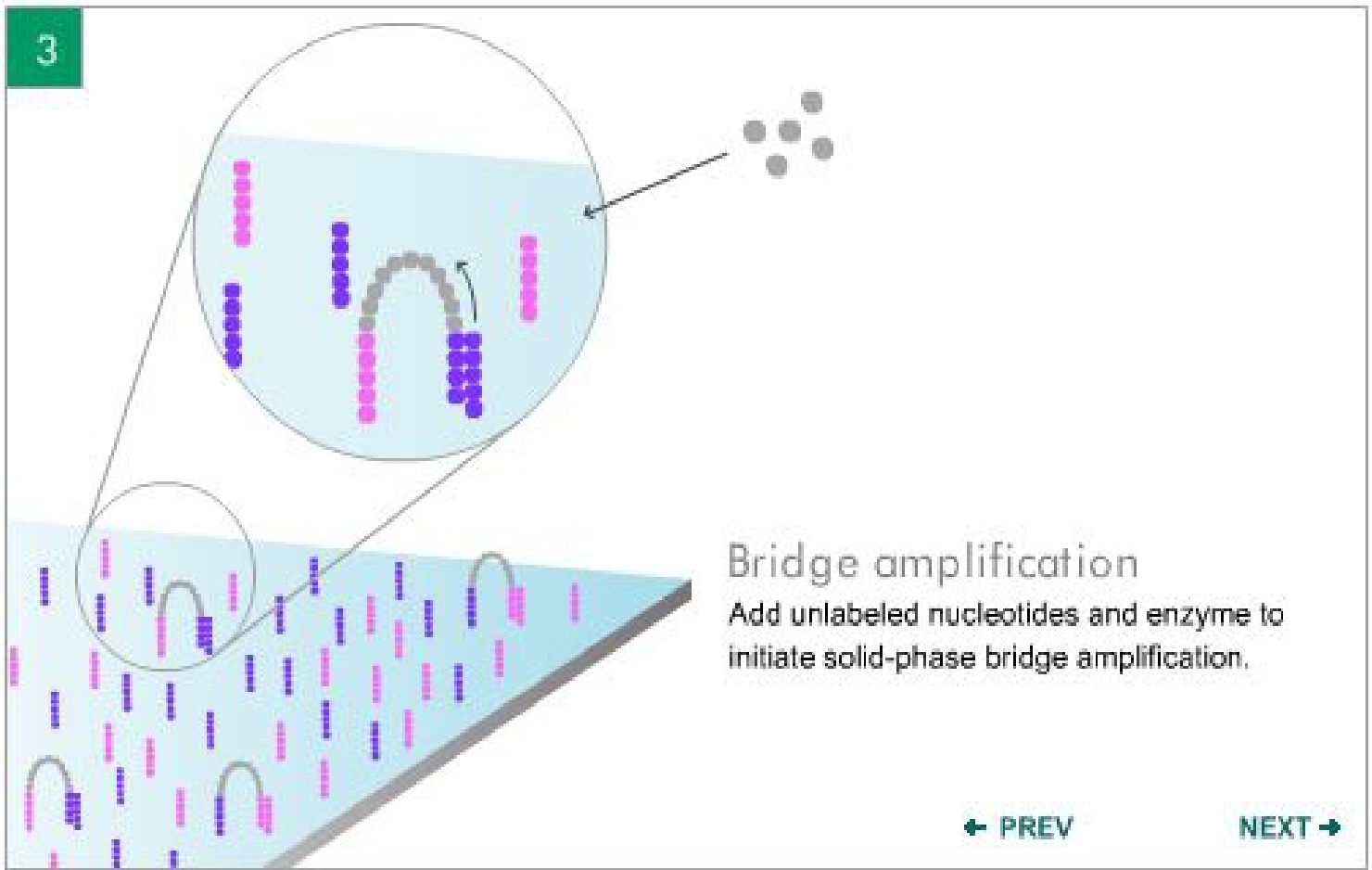
Sequencing-By-Synthesis Demo



Illumina Solexa Sequencing Technology

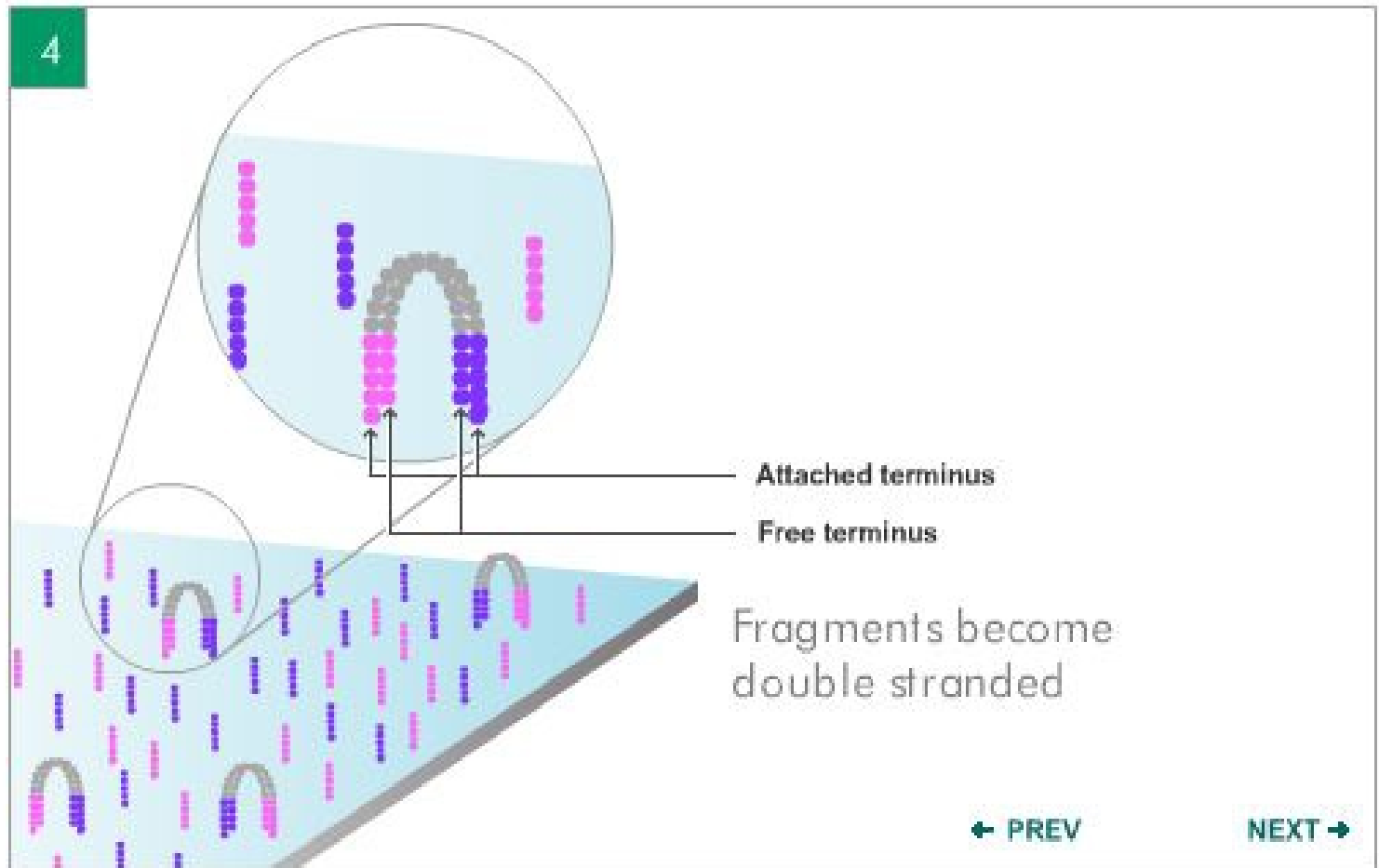


Sequencing-By-Synthesis Demo

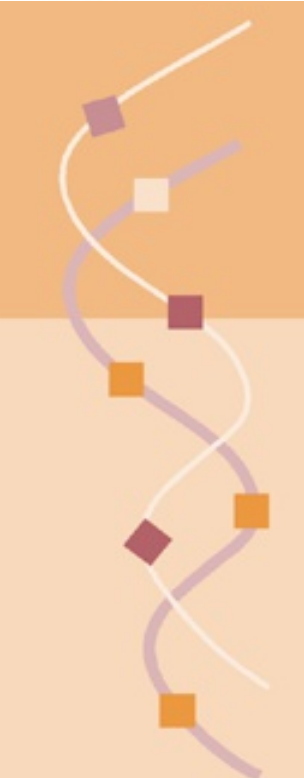


Illumina Solexa Sequencing Technology

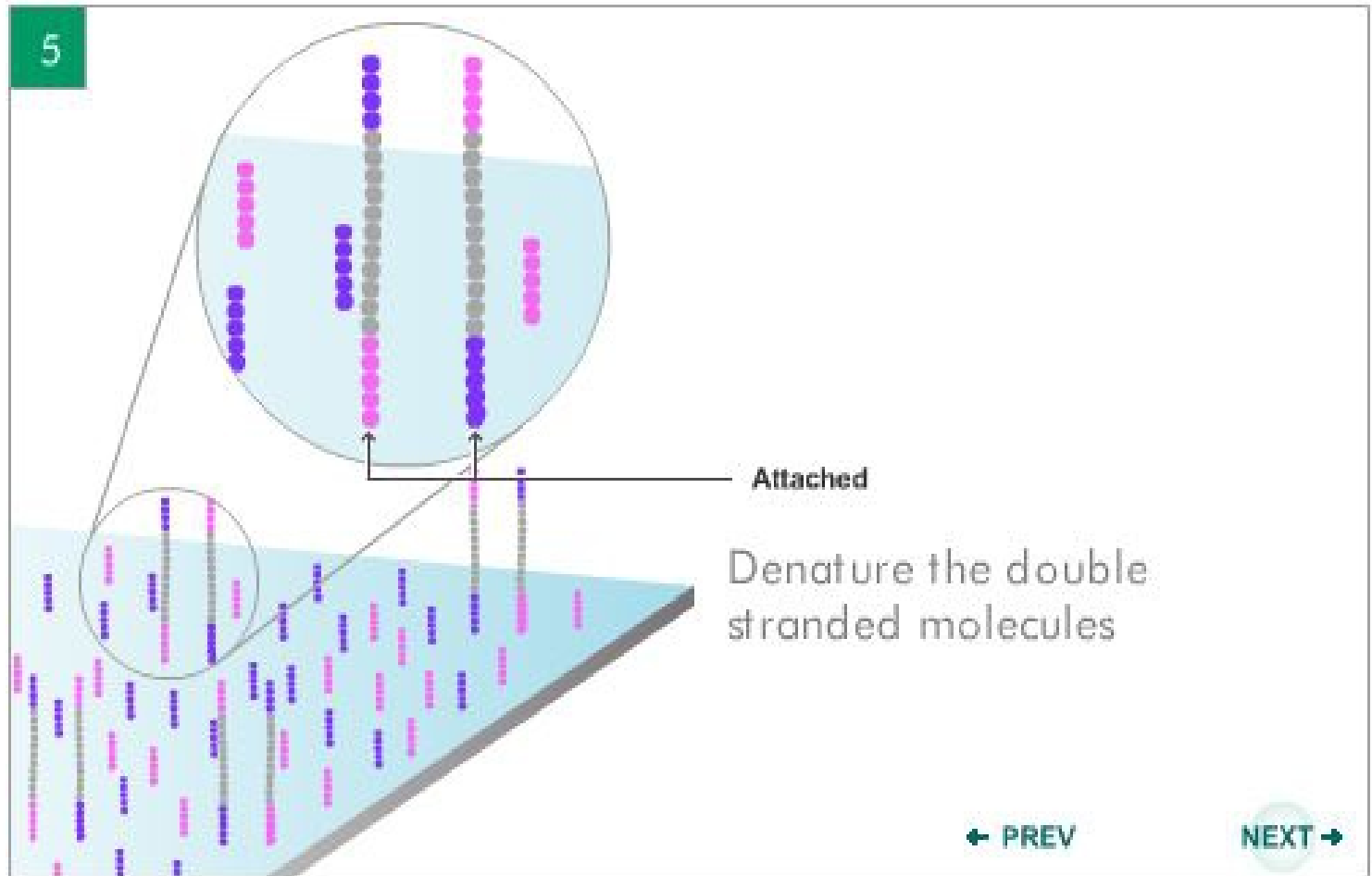
Sequencing-By-Synthesis Demo



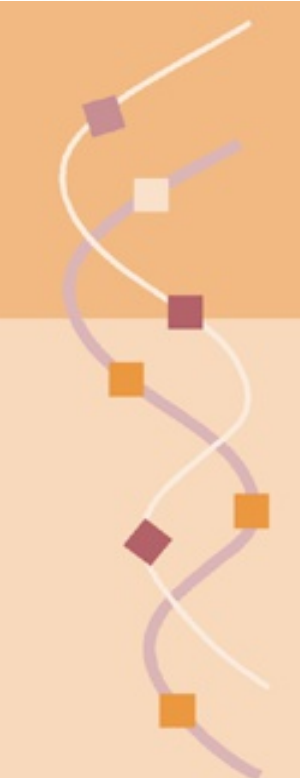
Illumina Solexa Sequencing Technology



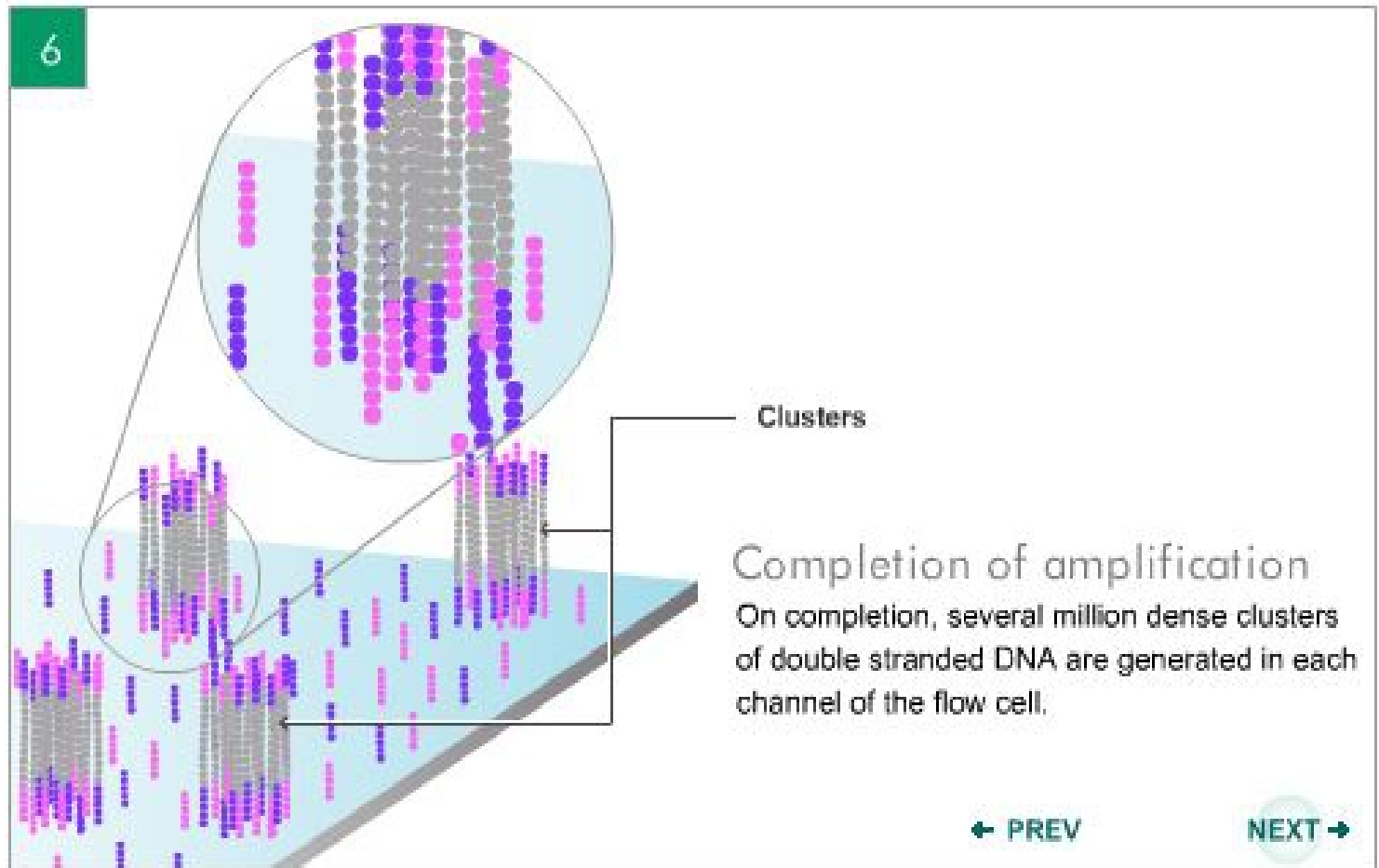
Sequencing-By-Synthesis Demo



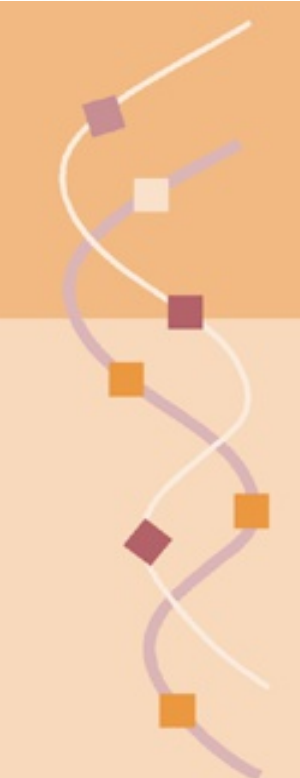
Illumina Solexa Sequencing Technology



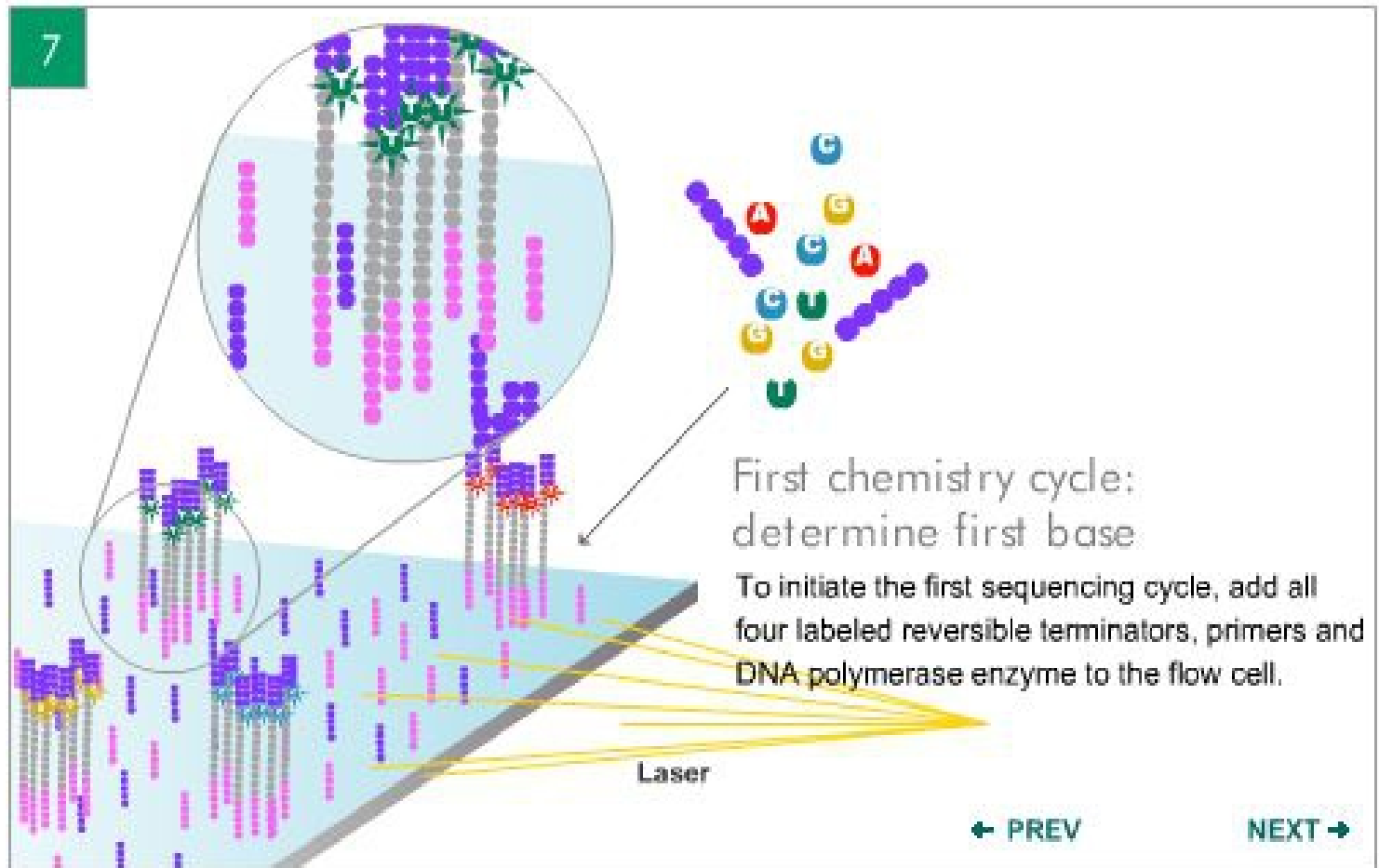
Sequencing-By-Synthesis Demo



Illumina Solexa Sequencing Technology



Sequencing-By-Synthesis Demo



Illumina Solexa Sequencing Technology



Sequencing-By-Synthesis Demo

8

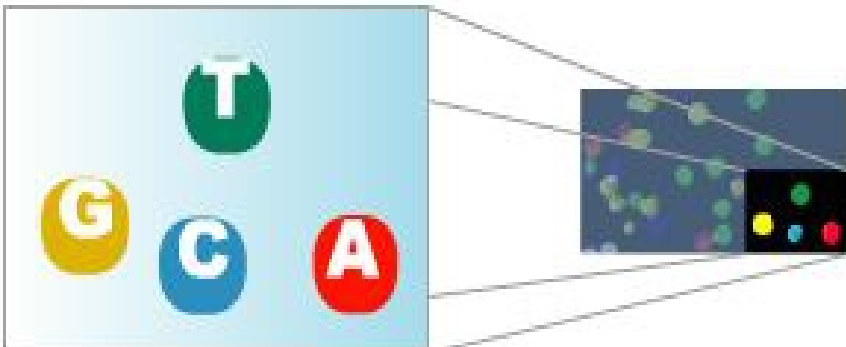


Image of first chemistry cycle

After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

Before initiating the next chemistry cycle

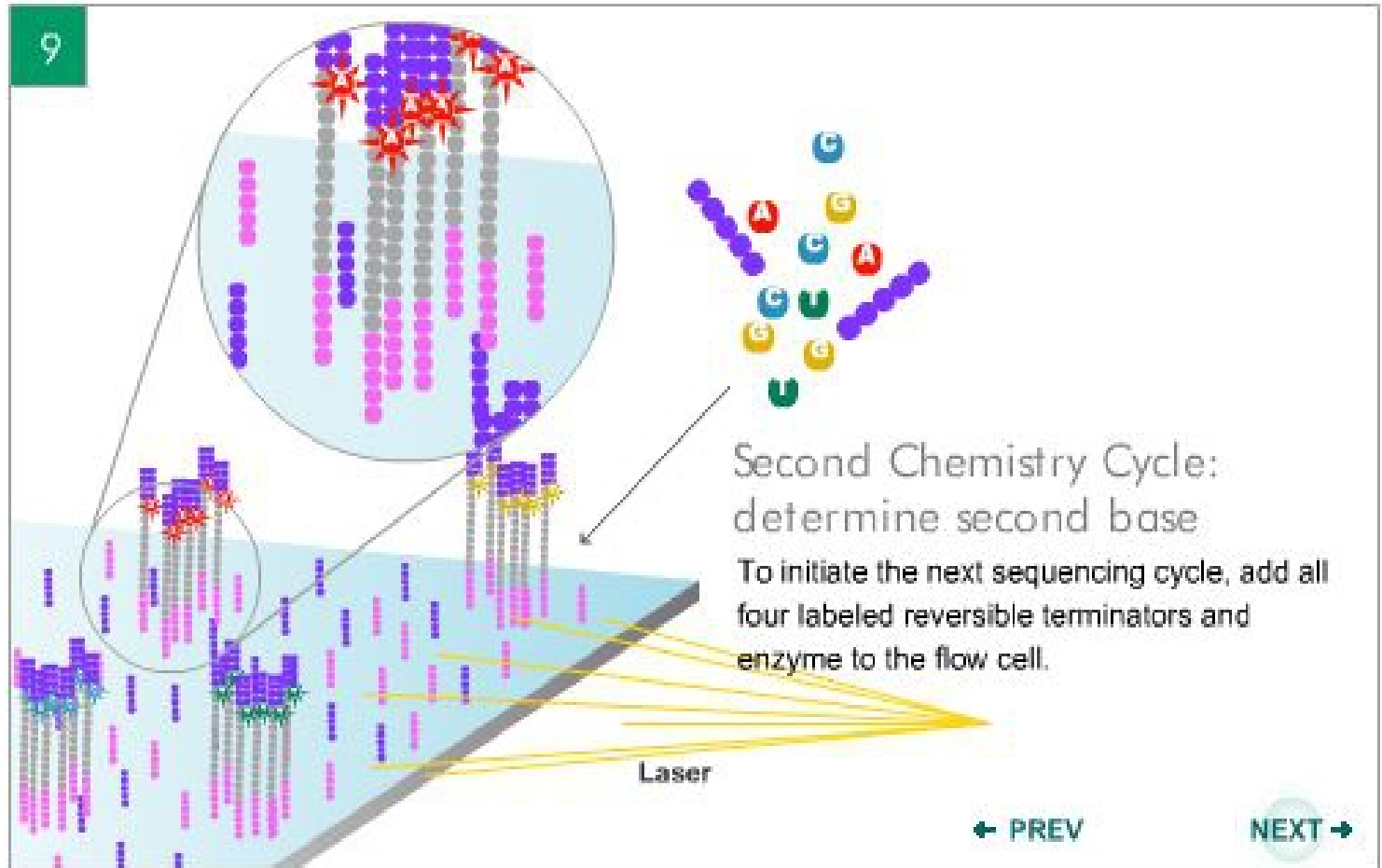
The blocked 3' terminus and the fluorophore from each incorporated base are removed.

← PREV
NEXT →

Illumina Solexa Sequencing Technology



Sequencing-By-Synthesis Demo



Illumina Solexa Sequencing Technology



Sequencing-By-Synthesis Demo

10


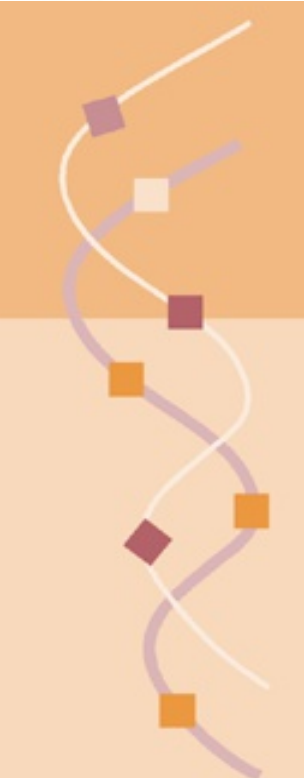


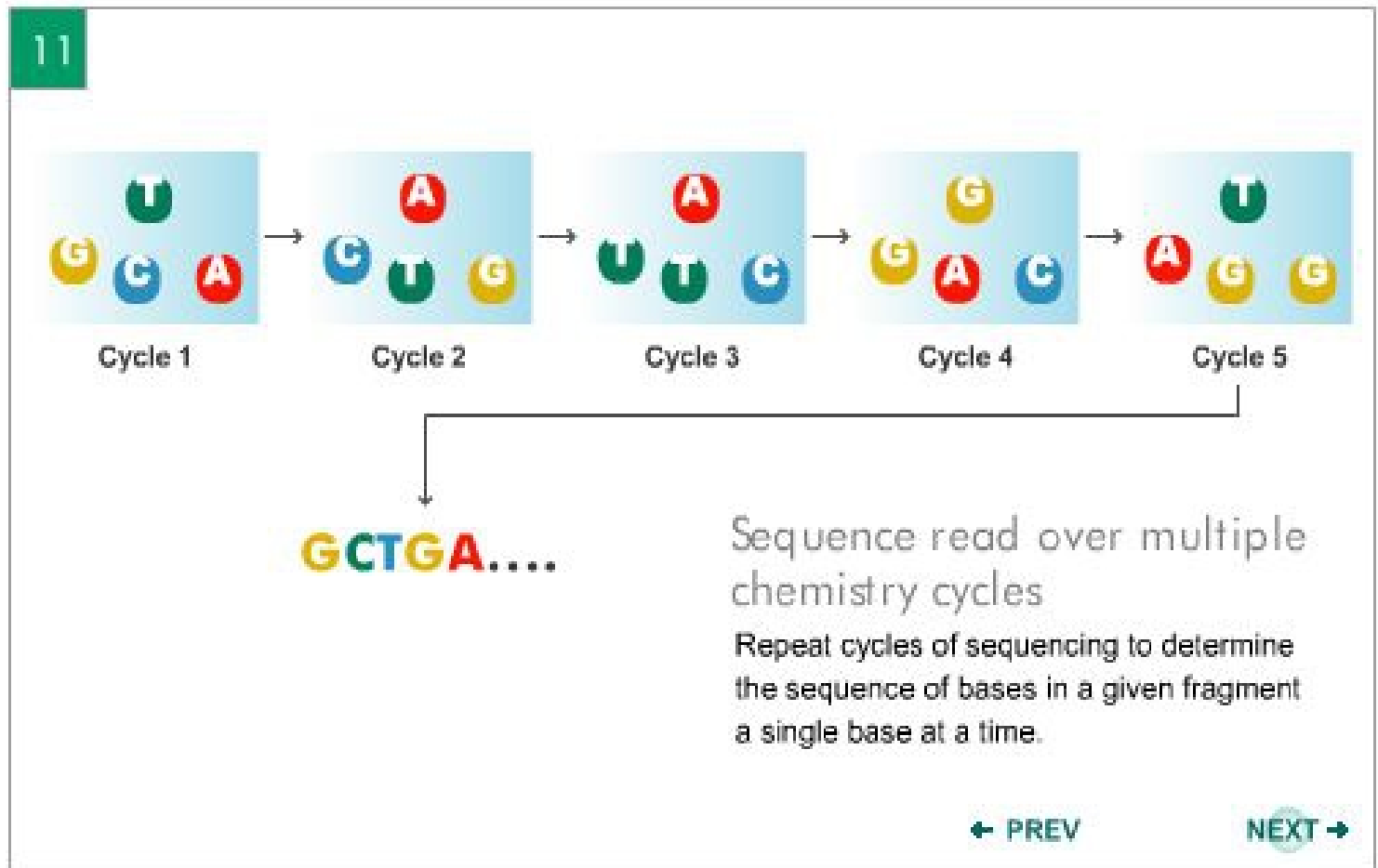
Image of second chemistry cycle is captured by the instrument
 After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

← PREV
NEXT →

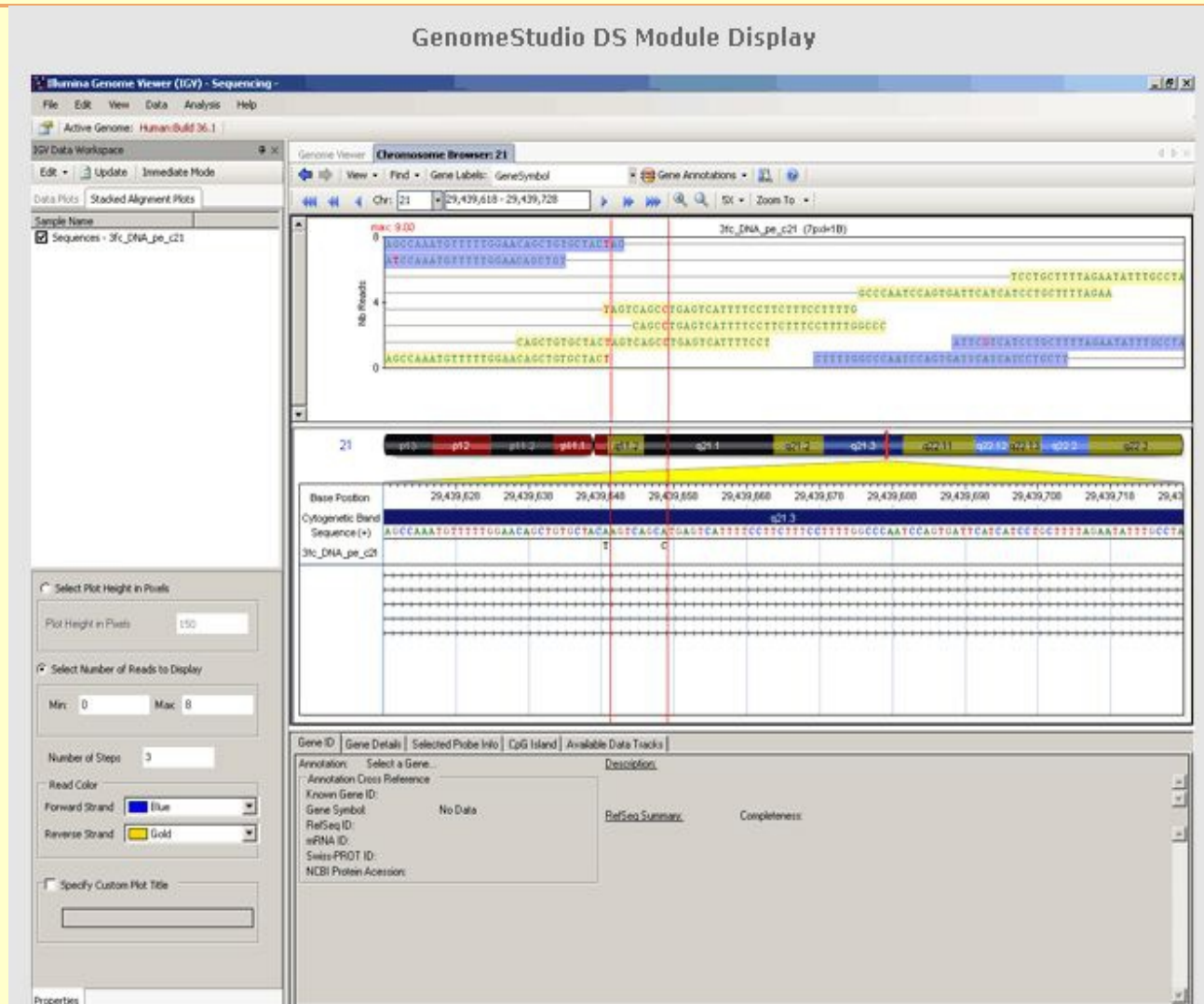
Illumina Solexa Sequencing Technology



Sequencing-By-Synthesis Demo



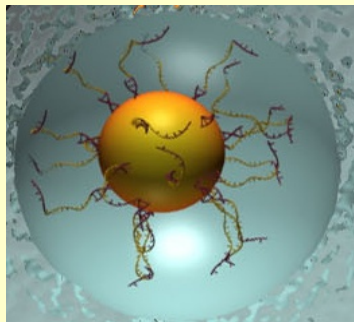
Illumina Solexa Sequencing Technology



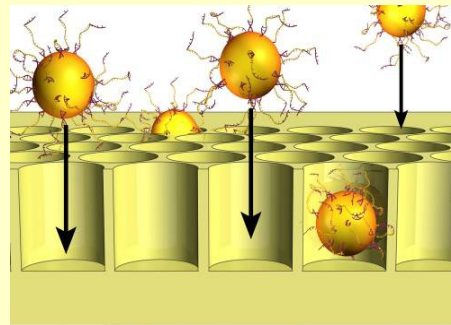
GenomeStudio displays SNPs identified by CASAVA based on alignment of reads against a reference sequence using the Illumina Chromosome Browser.

Life Sciences 454 Process Overview

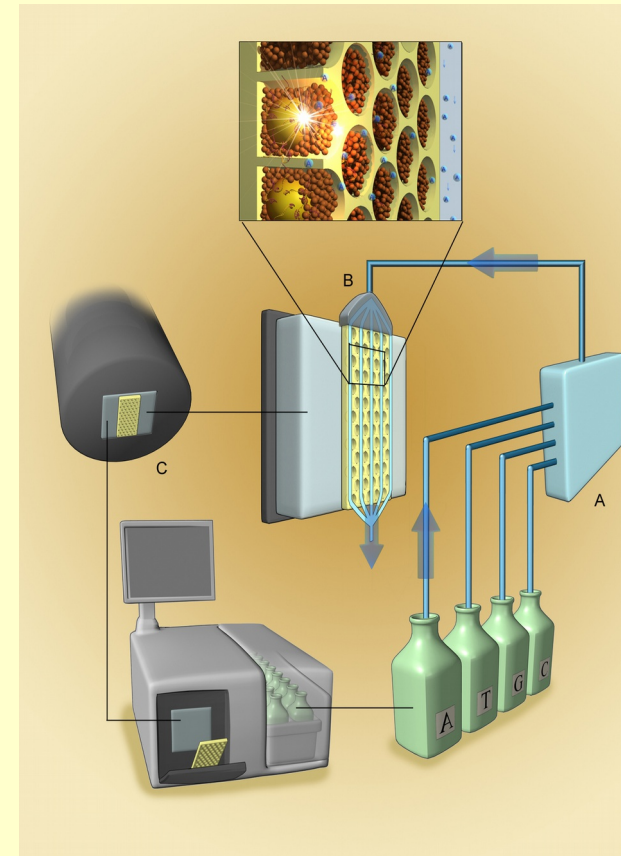
1) Prepare Adapter Ligated ssDNA Library



2) Clonal Amplification on 28 μ beads



3) Load beads and enzymes in PicoTiter Plate™




4) Perform Sequencing by synthesis on the 454 Instrument

Pacific Biosciences SMRT Sequencing

New PacBio Sequencing Technology Video

<http://www.pacificbiosciences.com/products/smrt-technology/>



PACIFIC BIOSCIENCES® APPLICATIONS **PRODUCTS** COMMUNITY SERVICES & SUPPORT NEWS & EVENTS ABOUT US
BLOG

Customers DevNet SampleNet Google Custom Search

Home | Products | SMRT Technology

SMRT TECHNOLOGY

Our SMRT technology harnesses the natural process of DNA replication, which is a highly efficient and accurate process. The enzyme responsible for replicating DNA in nature is called the DNA polymerase. The DNA polymerase attaches itself to a strand of DNA to be replicated, examines the individual base at the point it is attached, and then determines which of four building blocks, or nucleotides, is required to replicate that individual base. After determining which nucleotide is required, the polymerase incorporates that nucleotide into the growing strand that is being produced. After incorporation, the enzyme advances to the next base to be replicated and the process is repeated. Our SMRT technology enables the observation of DNA synthesis as it occurs in real time. To overcome the challenges inherent in observing an enzyme that is 15 nanometers, or nm, in diameter running in real time, we developed three key innovations:

- The SMRT Cell
- Phospholinked nucleotides
- The PacBio RS II

Watch an overview of SMRT Technology below.

PRODUCTS

- PACBIO RS II ›
Workflow ›

CONSUMABLES ›

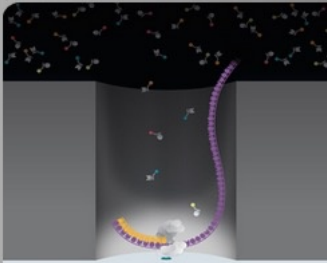
- SMRT Cells ›
- Reagents ›
- Disposables ›

SOFTWARE ›

- Instrument ›
- Analysis ›
- Algorithms ›


SMRT TECHNOLOGY ›

- SMRT Sequencing Advantage ›



With an active polymerase immobilized at the bottom of each ZMW, nucleotides diffuse into the ZMW chamber. In order to detect incorporation events and identify the base, each of the four nucleotides A, C, G and T are labeled with a different fluorescent dye having a distinct emission spectrum. Since the excitation illumination is directed to the bottom of the ZMW, nucleotides held by the polymerase prior to incorporation emit an extended signal that identifies the base being incorporated.

PLAYLIST | 27 / 52 Overview of SMRT Technology



SMRT® Cell

0:01 / 2:08

Pacific Biosciences SMRT Sequencing

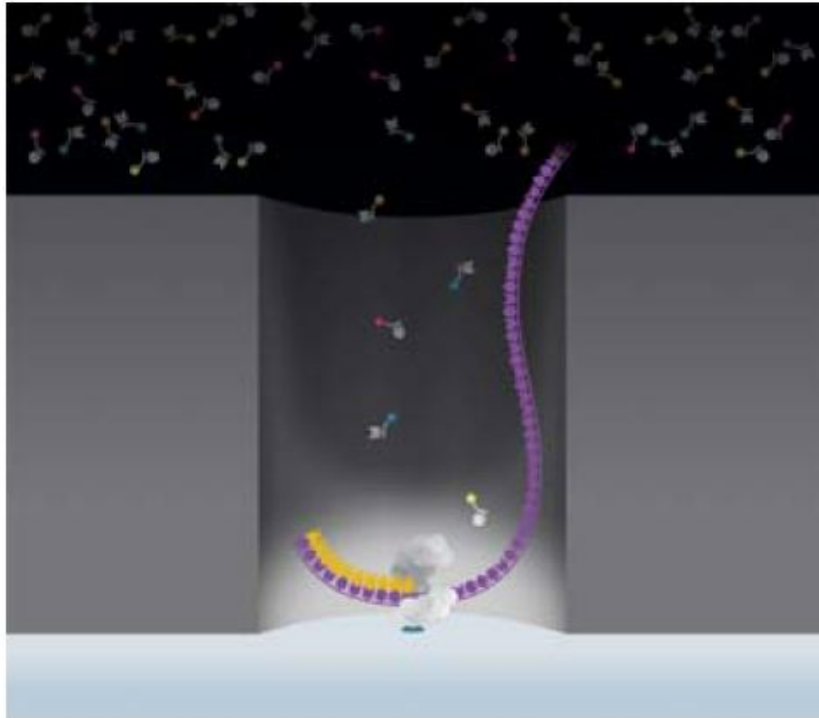


Figure 6. ZMW with DNA polymerase and phospholinked nucleotides

Phospholinked nucleotides are added into the ZMW at the high concentrations required for proper enzyme functioning.

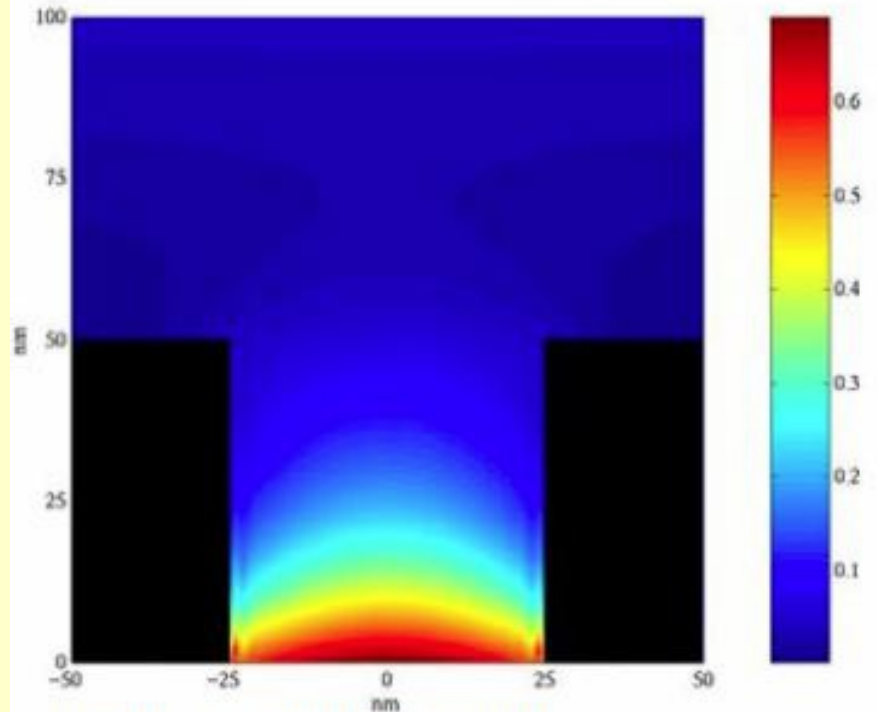


Figure 4. Detection volume

Attenuated light from the excitation beam penetrates only the lower 20-30 nm of each waveguide, creating a detection volume of 20 zeptoliters (10^{-21} liters).

Phospholinked Fluorophores

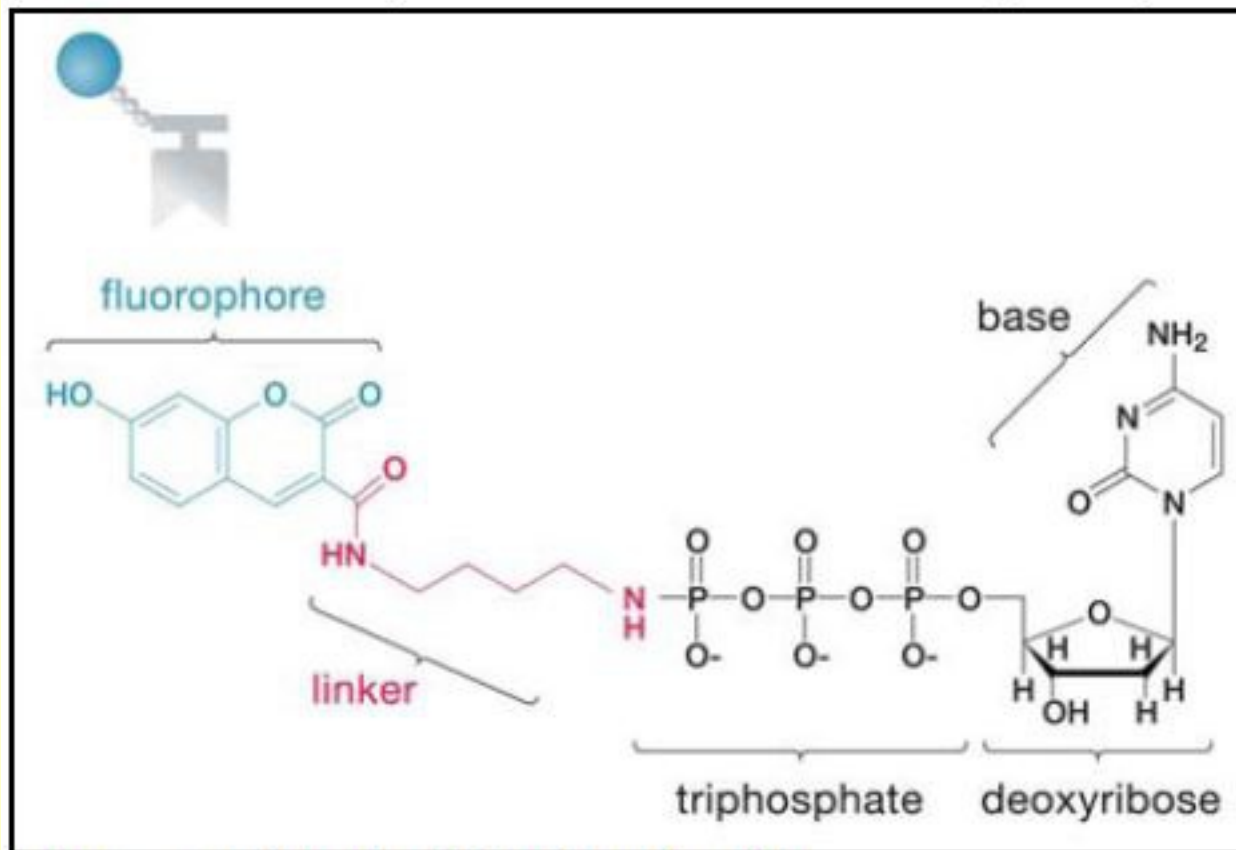
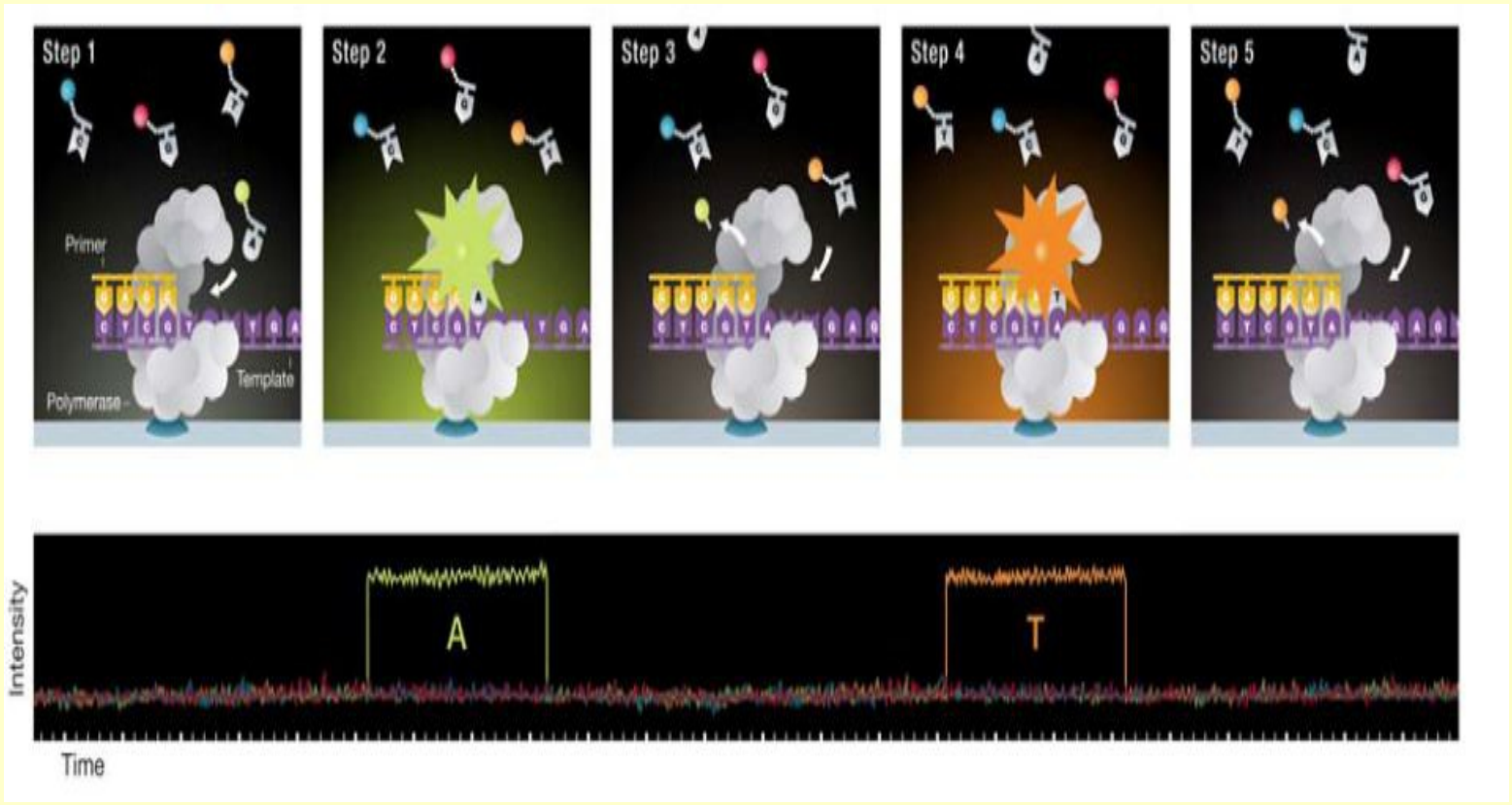


Figure 9. Phospholinked nucleotides

Phospholinked nucleotides have fluorophores attached to the triphosphate chain, which is naturally cleaved when the nucleotide is incorporated.

Processive Synthesis



Synthesis of Long Duplex DNA

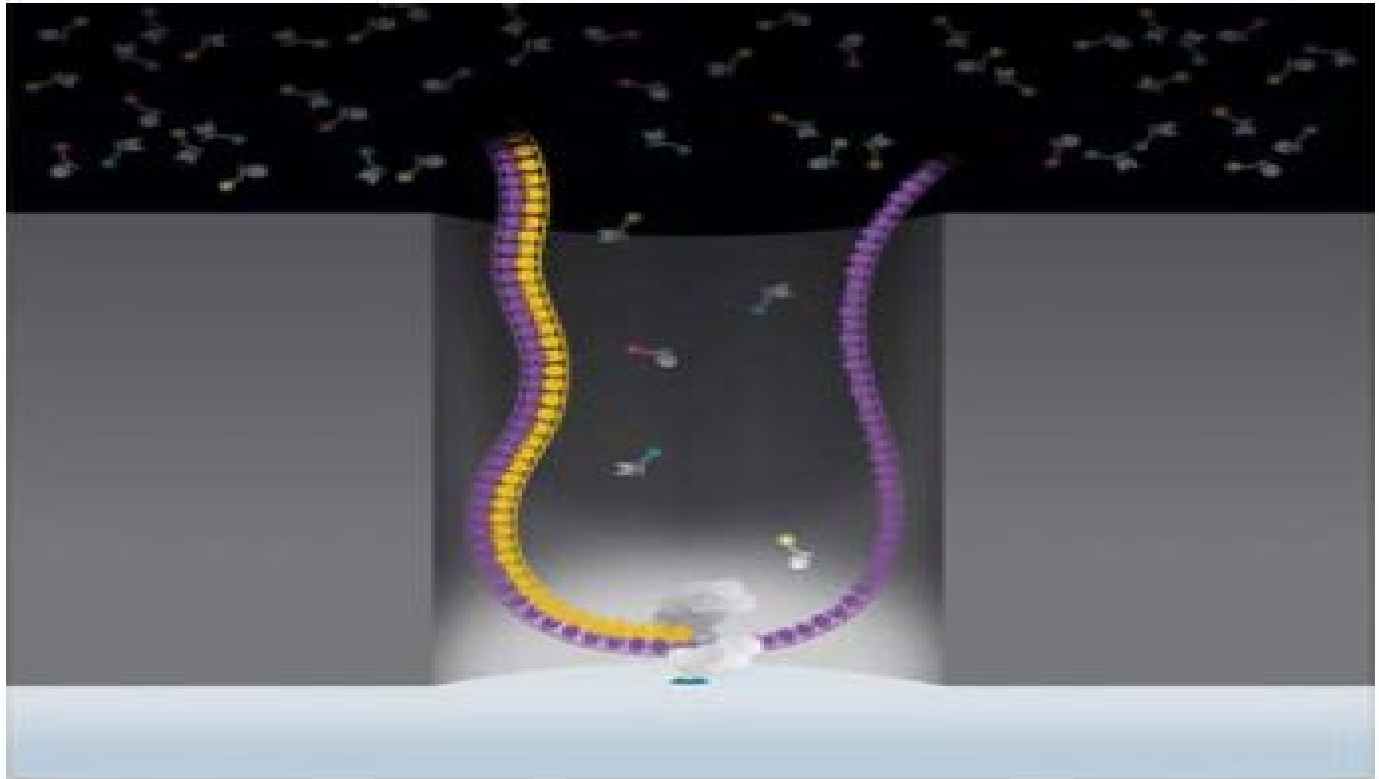
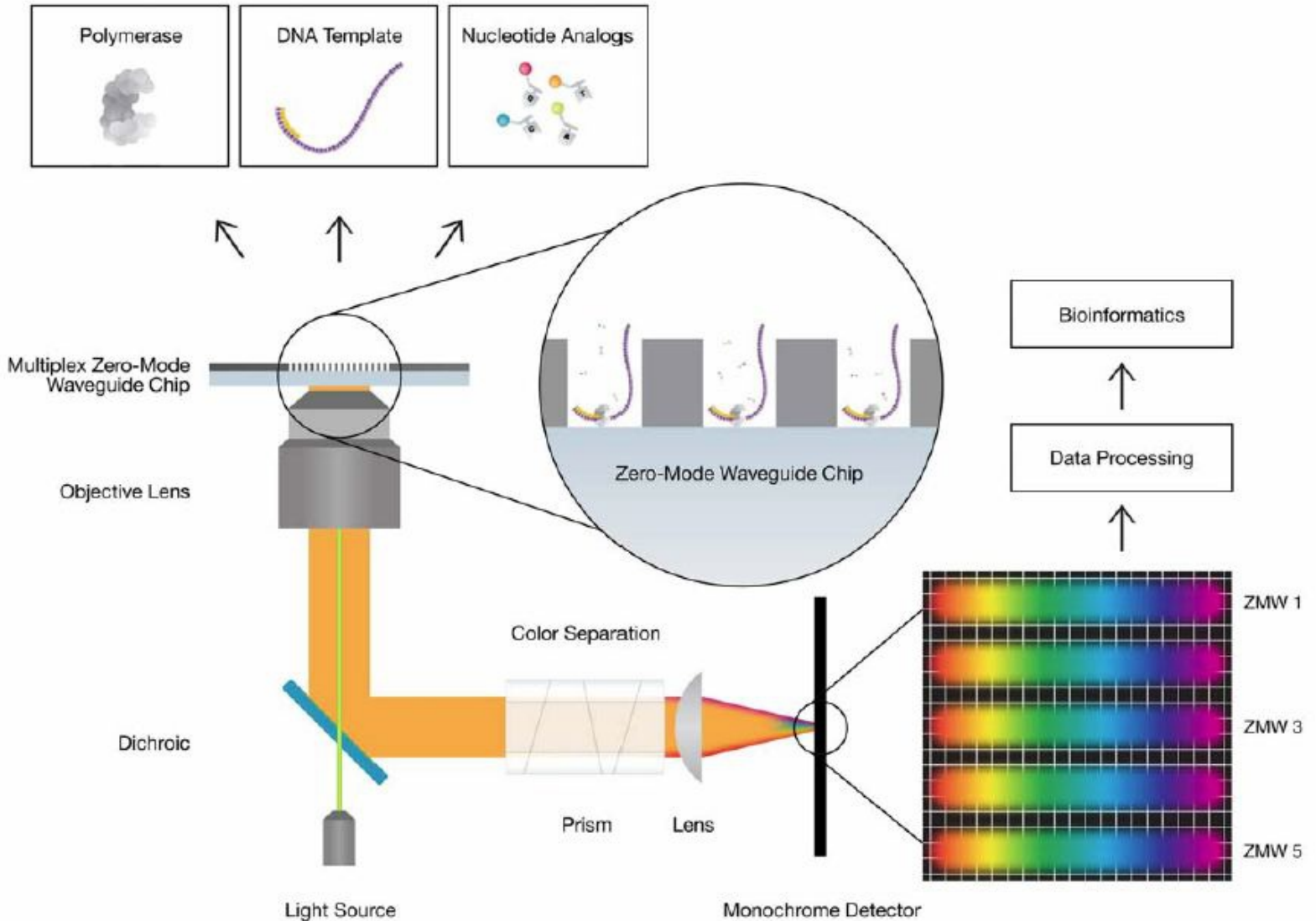


Figure 11. Synthesis of long DNA.

DNA polymerase processively incorporates nucleotides producing long, natural DNA.

Highly Parallel Optics System



Circular Templates Gives Redundant Sequencing and Accuracy



Circular Templates Gives Redundant Sequencing and Accuracy

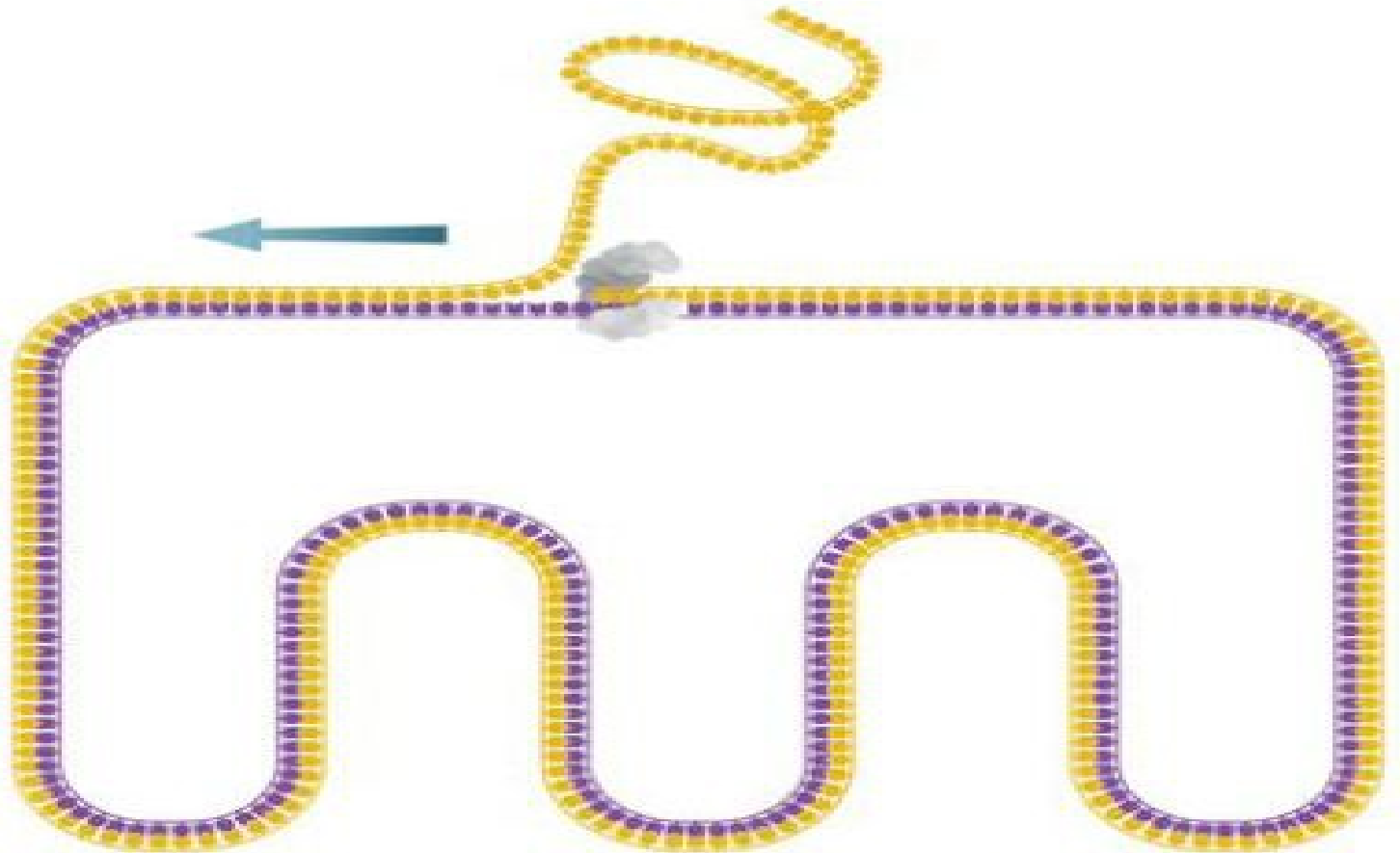


Figure 16. Molecular redundant sequencing

Ion Torrent Sequencing

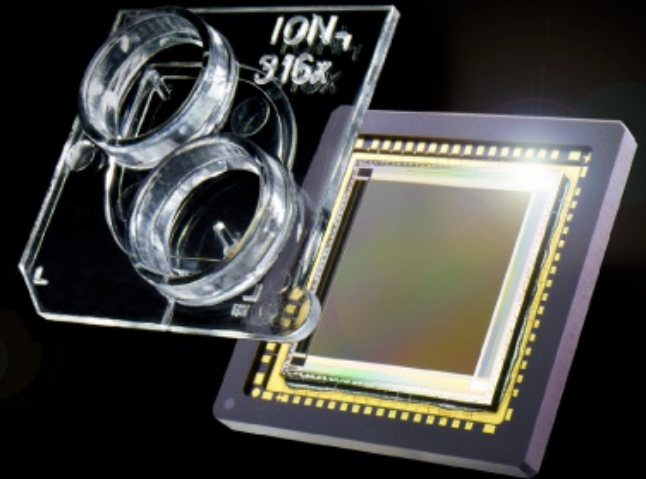


Semiconductor Sequencing for Life™



- HOME
- TECHNOLOGY
- PRODUCTS
- APPLICATIONS
- COMMUNITY
- SUPPORT
- STORE

fastest next-gen workflow
10X more throughput
fastest-selling sequencer
all in six months



Ion Torrent
Ion 316™ -- Everything moves faster
when The Chip is the Machine™

[Watch the video >](#)
[Read more >](#)

Publications

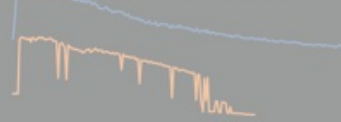
nature

An integrated semiconductor device enabling non-optical genome sequencing

Application Note

The Ion PGM™ sequencer exhibits superior long-read accuracy

[Learn more >](#)



Life Grand Challenges



News

August 2011
[Combating Superbug...more >](#)

[Ion Torrent long read accuracy, download app note & data...more >](#)

[View more Ion News...go >](#)

Ion Torrent Sequencing



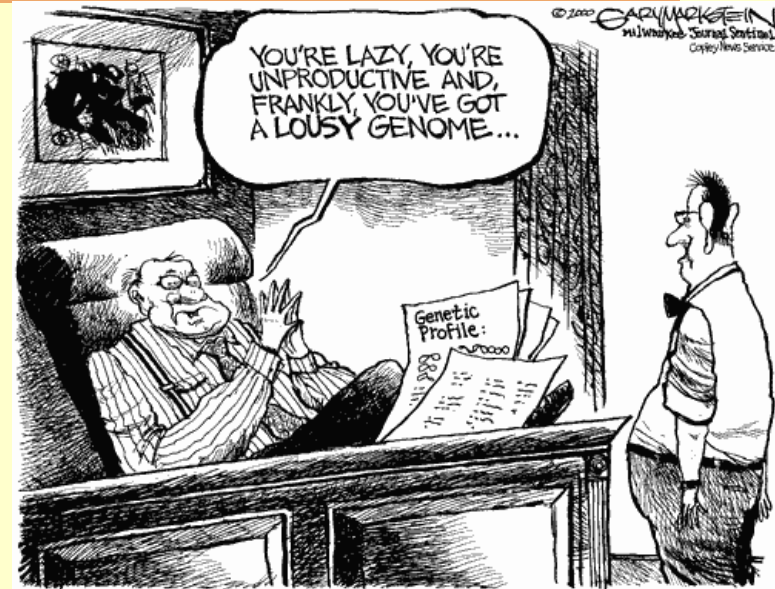
Ion Torrent Sequencing



The Human Genome

How fast is the cost going down?

- 2006: \$ 50 million
- 2008: \$500,000
- 2009: \$50,000
- 2010: \$20,000
- 2011: \$5,000
- 2012: \$4,000
- 2013: \$3,000
- 2014 \$1,700

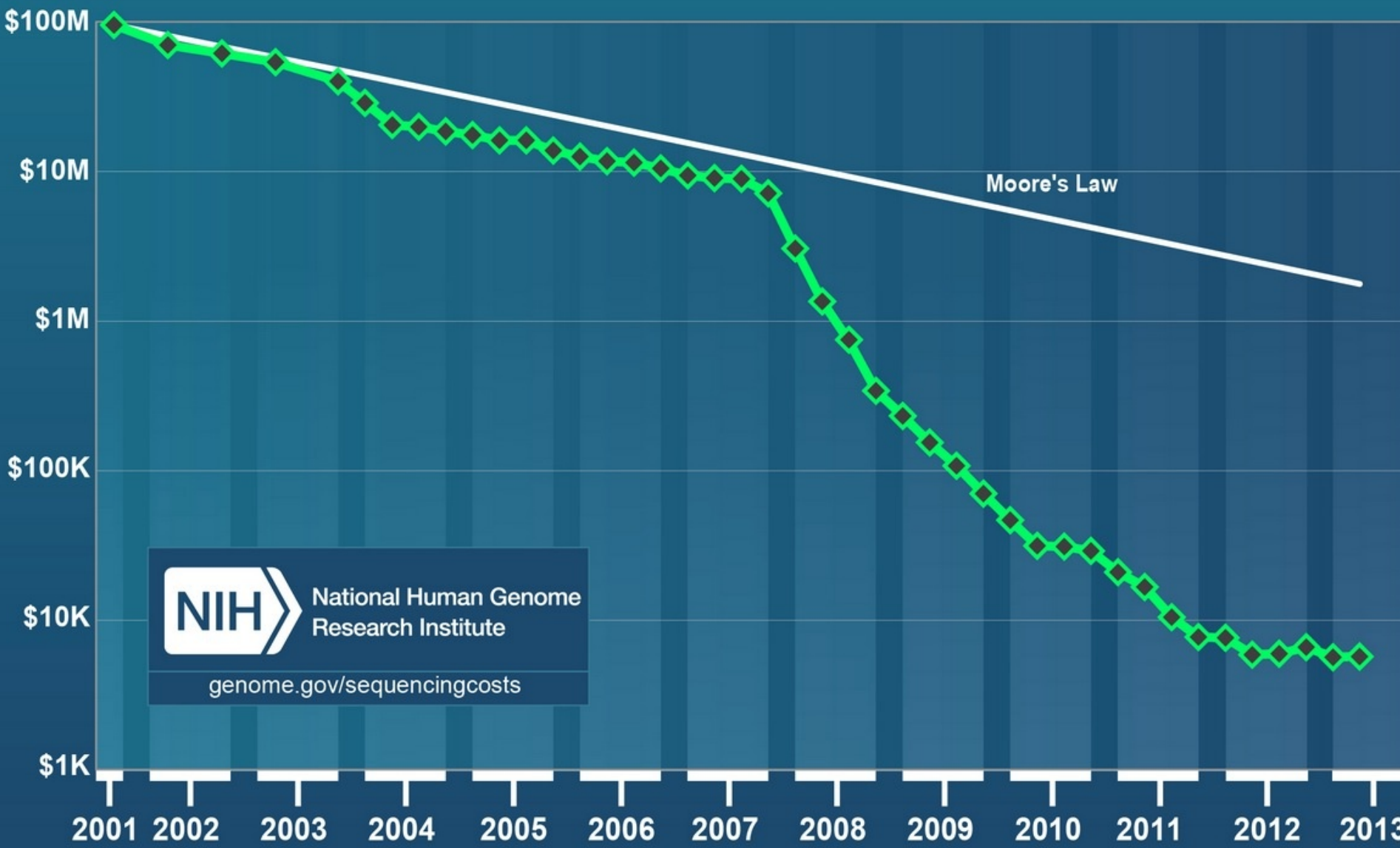


"The locket contains a strand of my DNA."



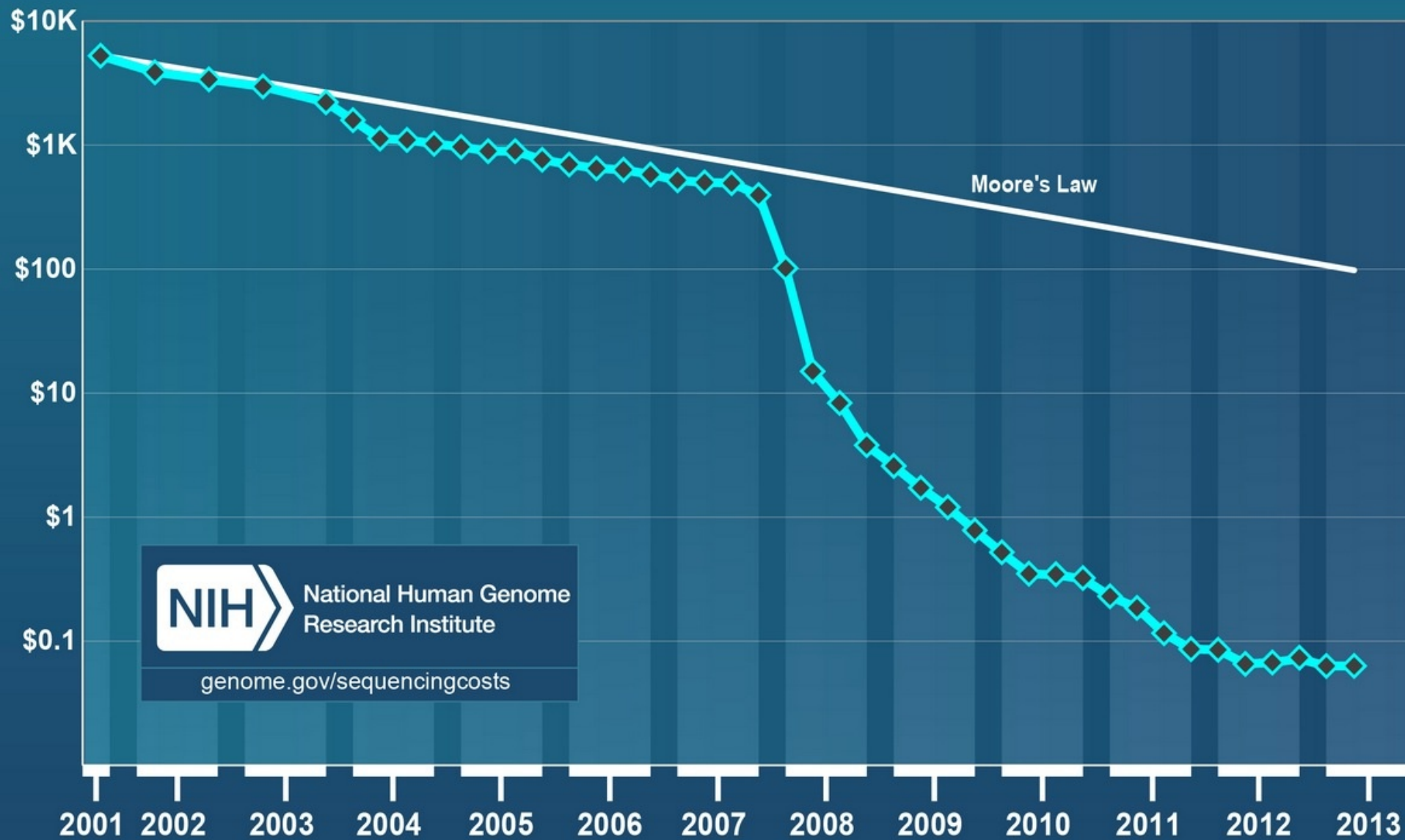
Thanks to Seraf in Batzoglou

Cost per Genome



NIH National Human Genome Research Institute
genome.gov/sequencingcosts

Cost per Raw Megabase of DNA Sequence



Components of a Typical Human Gene



Active Genes are Transcribed into RNA

Promoter

Exon

Intron

Exon

Intron

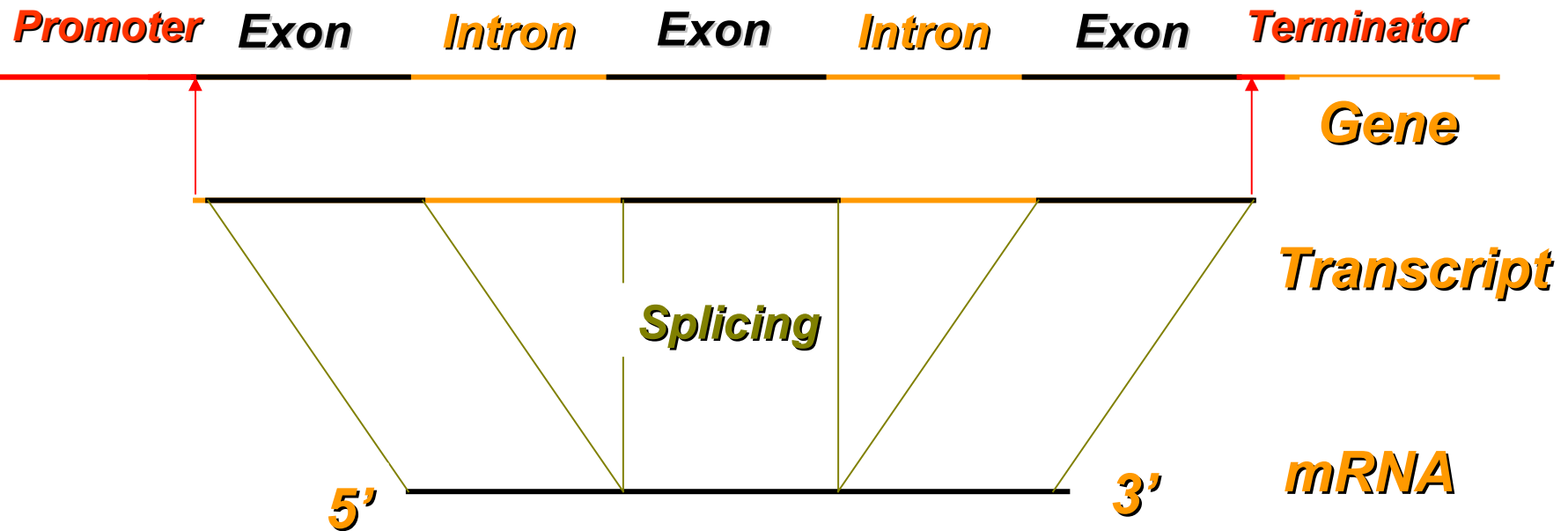
Exon

Terminator

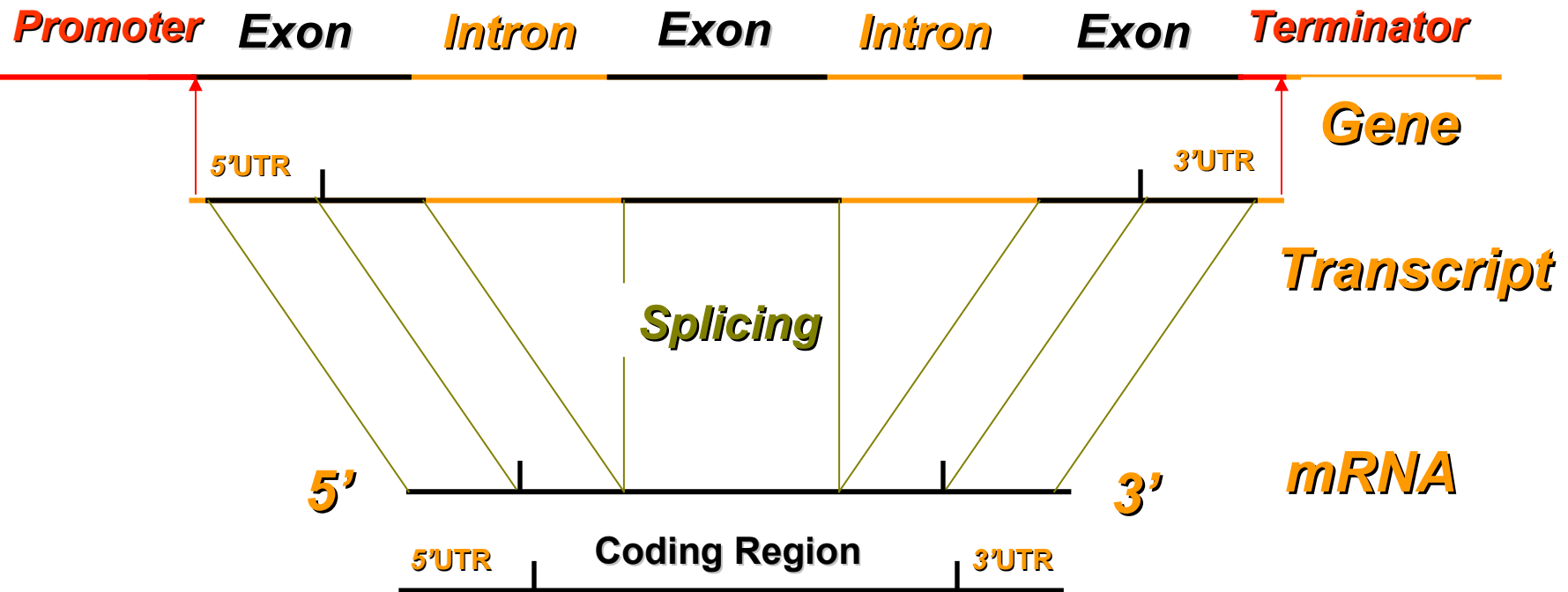
Gene

**Primary
Transcript**

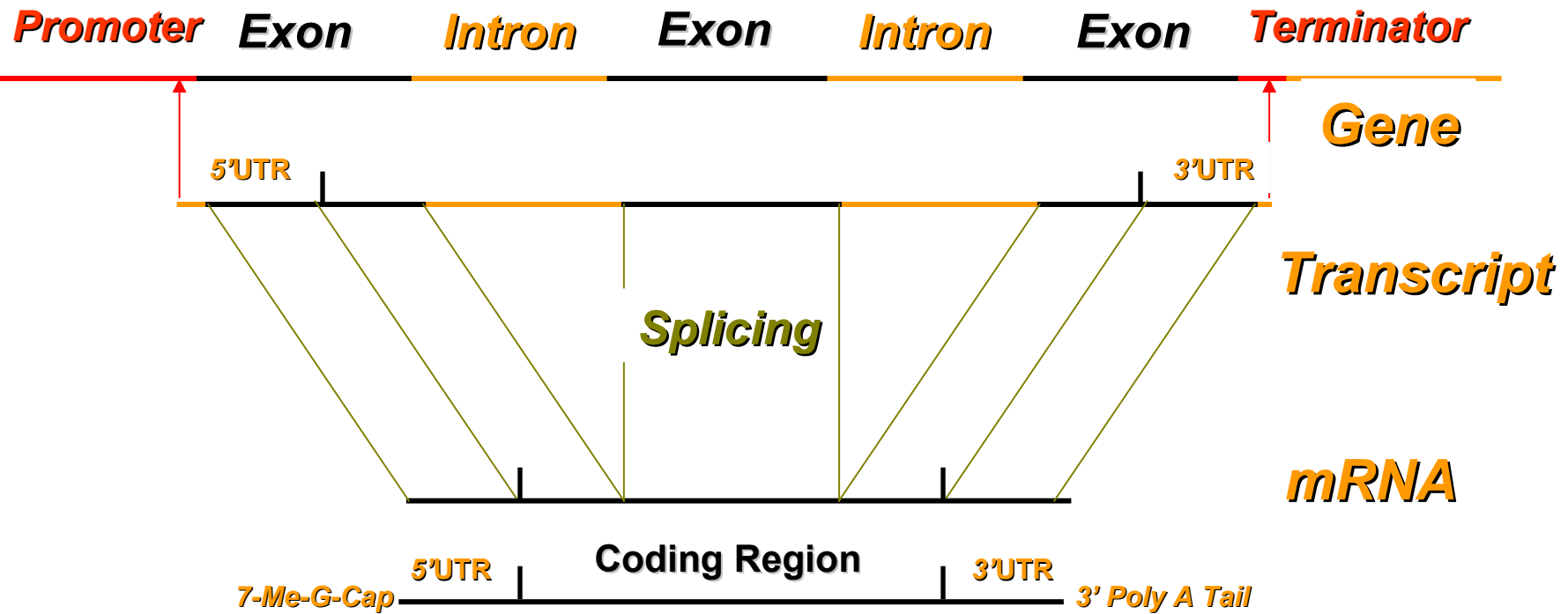
Splicing Transcript Yields Mature mRNA



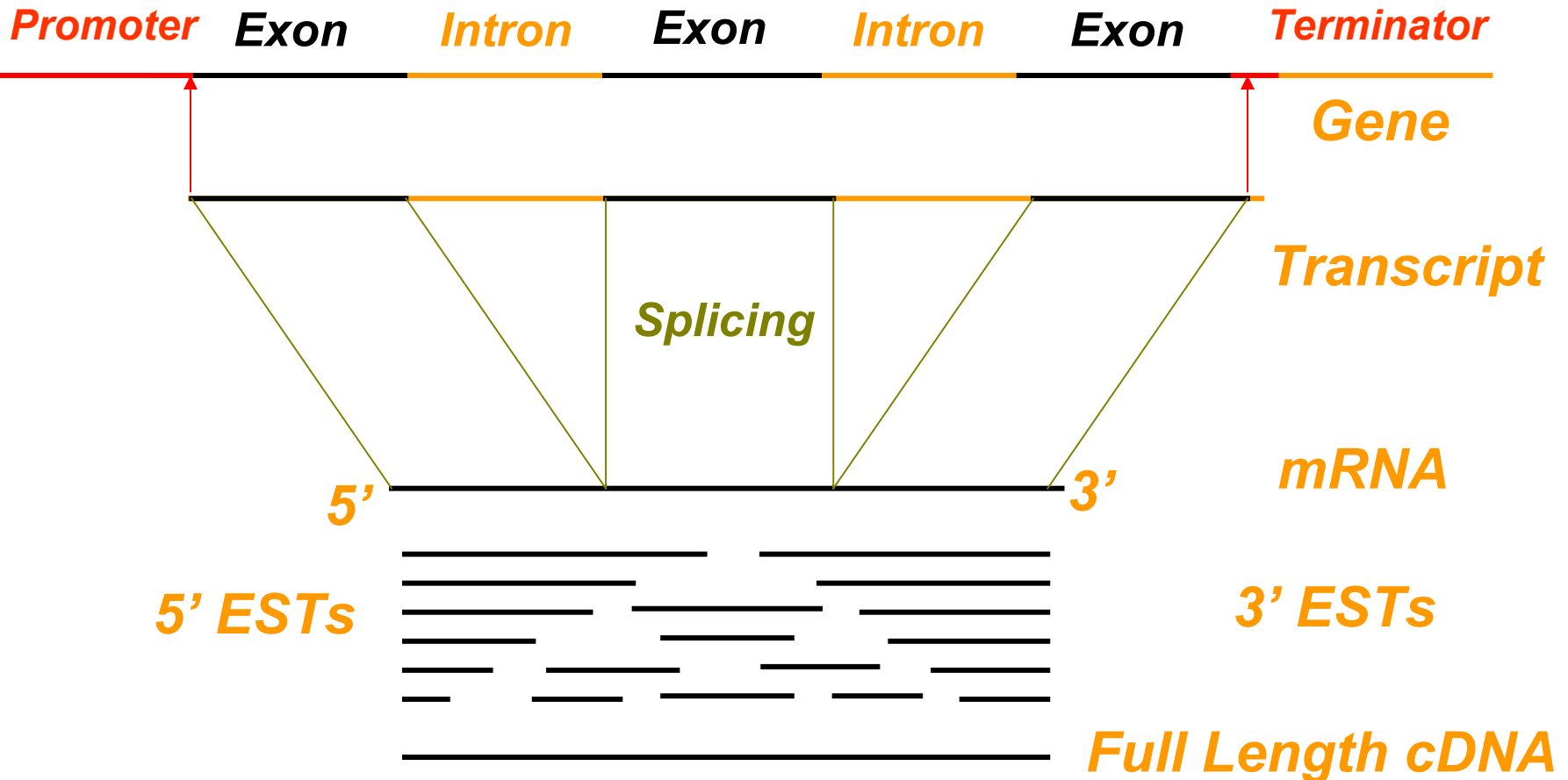
Mature mRNA contains Coding Region and 5' and 3' Untranslated Regions



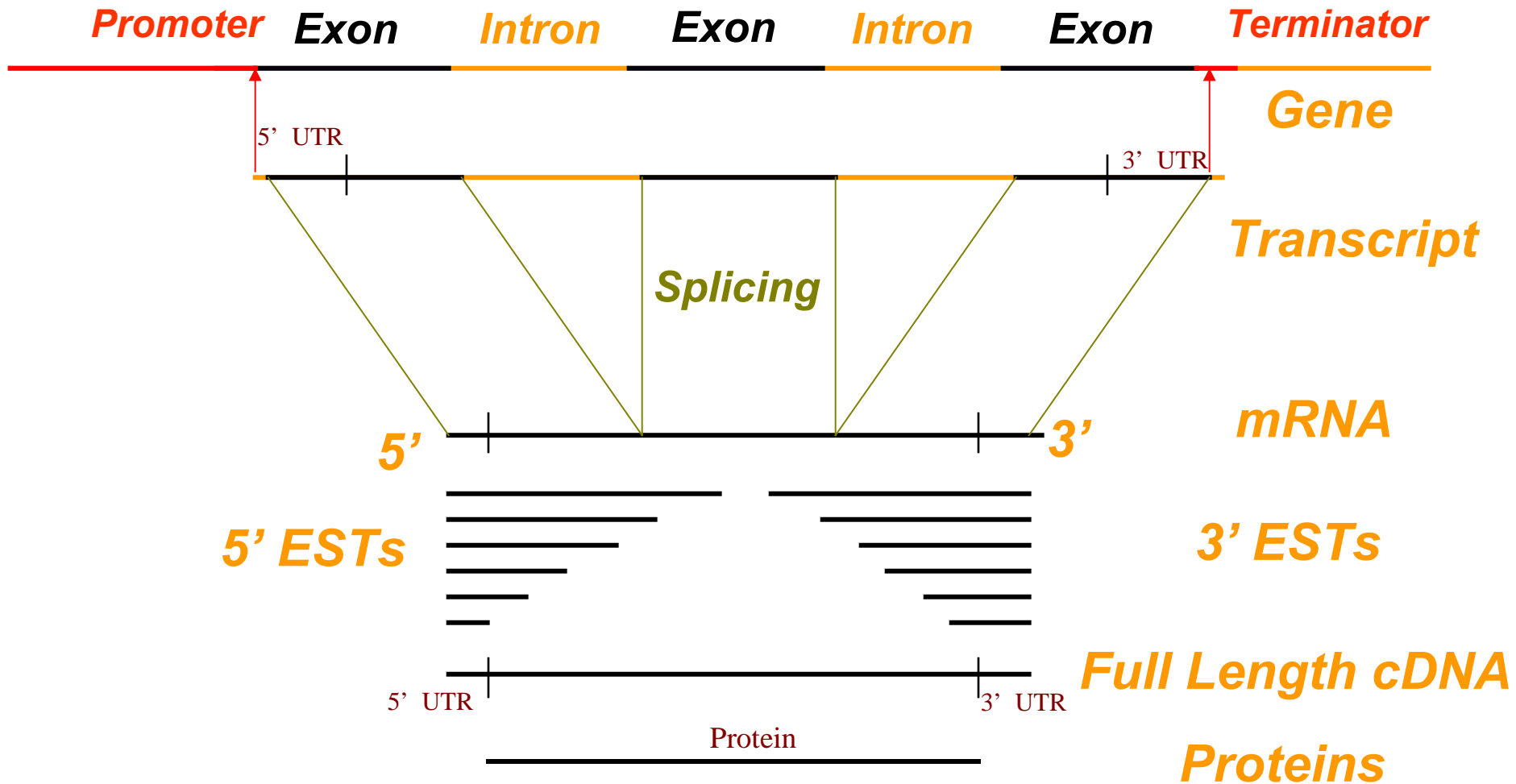
Mature mRNA contains 7-Methyl-Guanosine 5' Cap and 3' Poly A Tail



ESTs, Full Length cDNA UniGene & RefSeq Databases



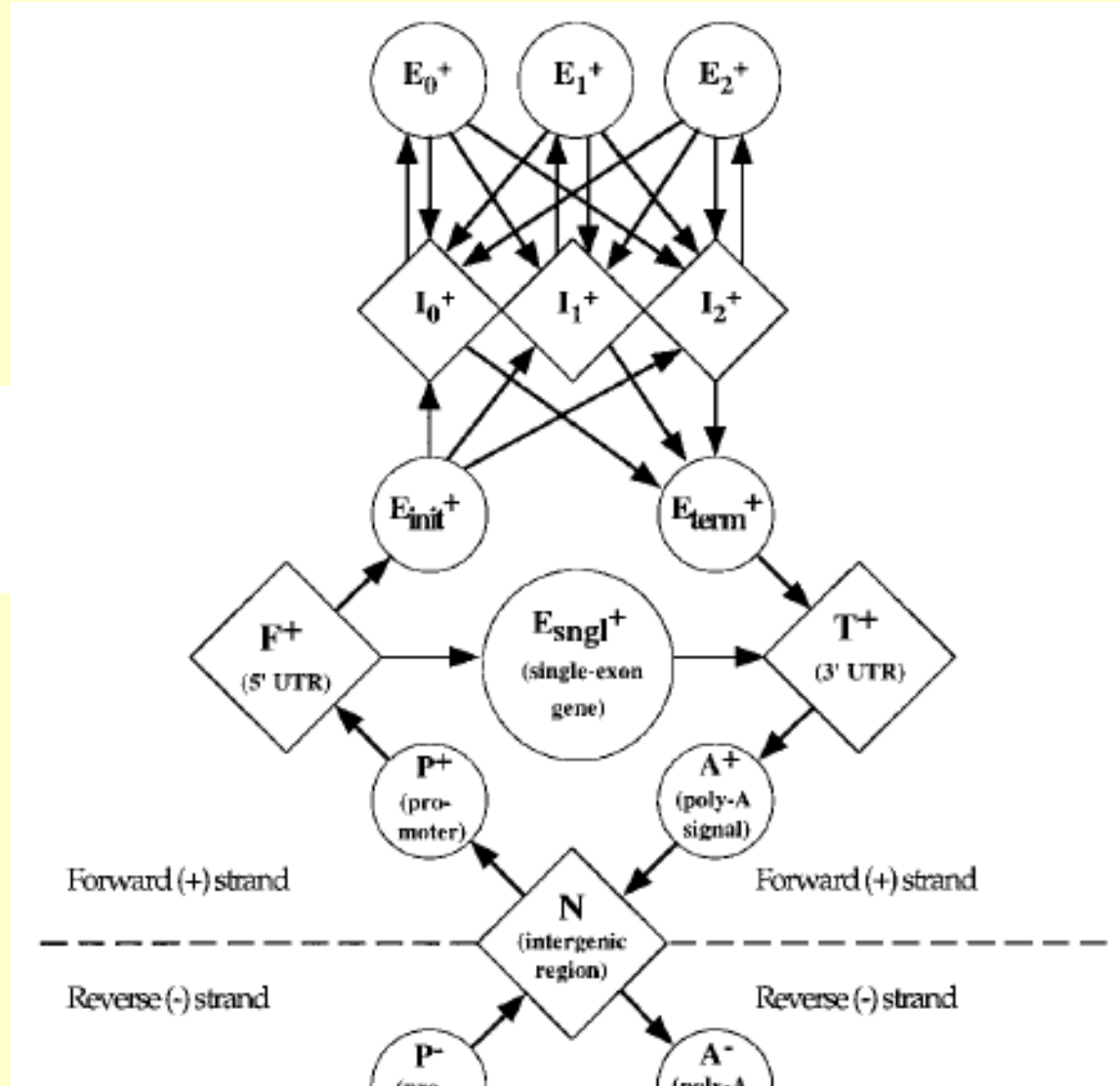
ESTs, Full Length cDNA UniGene & RefSeq Databases



GENSCAN Gene Model

<http://genes.mit.edu/GENSCAN.html>

Hidden Markov models of gene structure



Genome Databases

Assembled contigs

A Mapping

dbSTS

dbSNP

B Gene Prediction

GrailEXP

GenScan

FGENESH

FGENESH+

GeneMark

C Expression Data

Human ESTs

UniGene Human

RefSeq Human

Ensembl cDNA

Mouse ESTs

Entrez Gene Mouse

RefSeq Mouse

D Protein Similarity

nrPRO

pFAM Motifs

E Additional Data

Promoters

F Summary

Entrez Gene

UCSC Browser

Ensembl

Entrez Gene Loci

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

NR-Pro

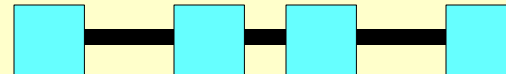


**Inclusive
Exon
Prediction**

UniGene



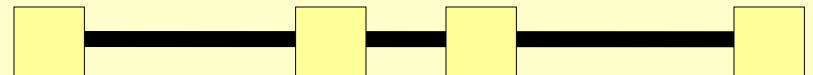
ESTs



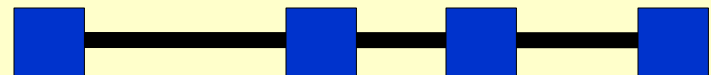
GrailEXP



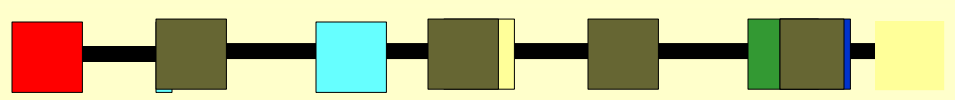
FGENESH



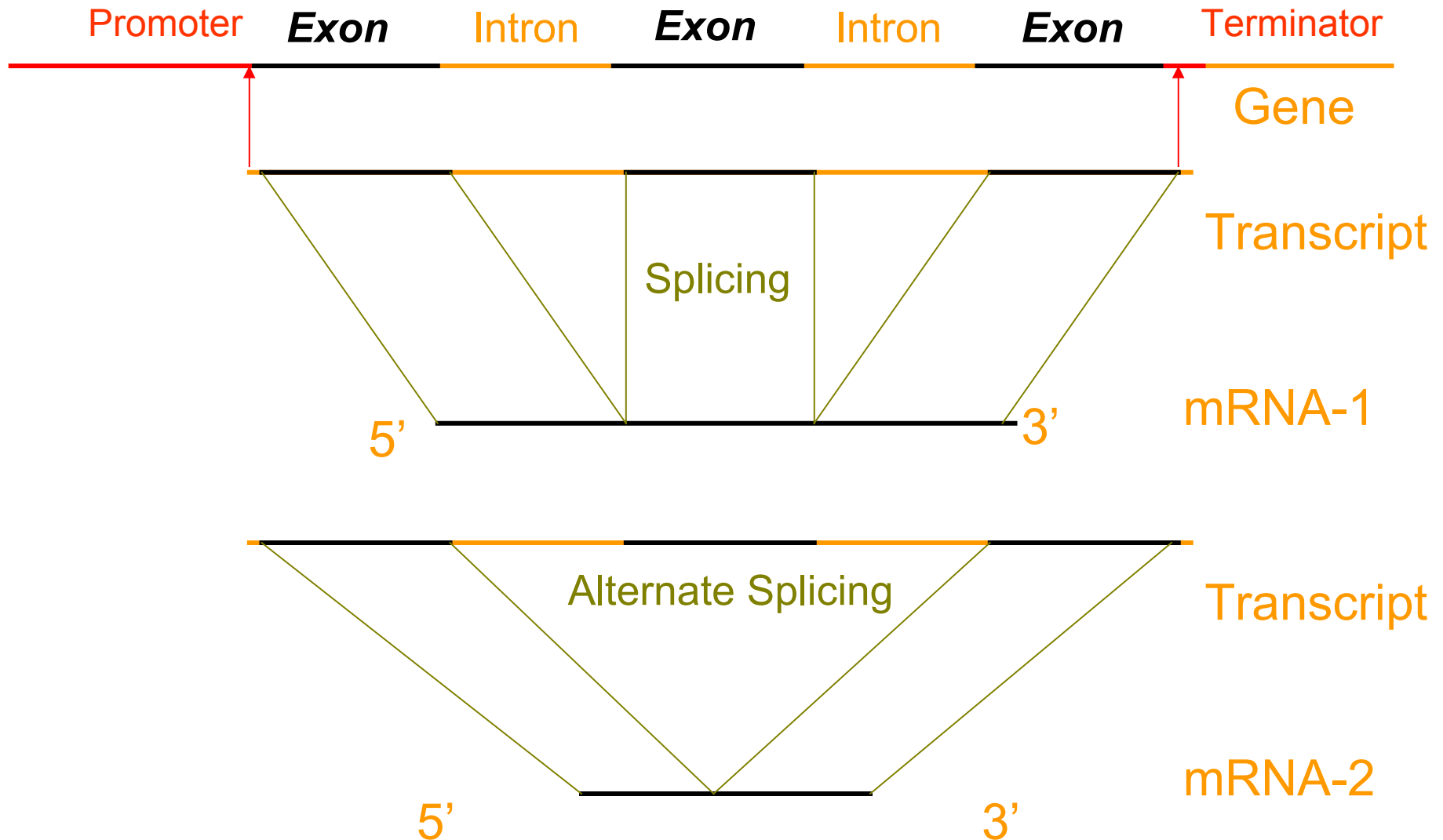
Genscan



Entrez Gene



Alternative Splicing Generates Distinct Proteins in Different Tissues



Genome

Genome ▾

dog



Search

[Limits](#) [Advanced](#)

[Help](#)



Genome

This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.

Using Genome

[Help](#)

[Browse by Organism](#)

[Download / FTP](#)

[Submit a genome](#)

Custom resources

[Human Genome](#)

[Microbes](#)

[Organelles](#)

[Plants](#)

[Viruses](#)

Other Resources

[Assembly](#)

[BioProject](#)

[BioSample](#)

[Map Viewer](#)

[Protein Clusters](#)

Genome Tools

[BLAST the Human Genome](#)

[Genomic groups BLAST](#)

[NCBI remap](#)

[Genome Decoration Page](#)

Genome Annotation and Analysis

[Eukaryotic Genome Annotation](#)

[Prokaryotic Genome Annotation](#)

[PASC \(Pairwise Sequence Comparison\)](#)

[TaxPlot \(3-way Genome Comparison\)](#)

External Resources

[GOLD - Genomes Online Database](#)

[Ensembl Genome Browser](#)

[Bacteria Genomes at Sanger](#)

[Large-Scale Genome Sequencing \(NHGRI\)](#)


Eukaryote Genomes

<http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>



Canis lupus familiaris Genome

<http://www.ncbi.nlm.nih.gov/sites/entrez?db=bioproject&cmd=Retrieve&dopt=O>


[Resources](#) [How To](#)
brutlag My NCBI Sign Out

BioProject
[Limits](#) [Advanced](#) [Help](#)

[Display Settings](#)

[Send to](#)

Name: **Canis lupus familiaris (dog)** Accession: PRJNA10726 ID: 10726
Title: **Model organism that is notable for its extensive genetic diversity and morphological variation**

The dog, *Canis lupus familiaris*, is a useful model organism for medical research due to extensive genetic diversity and morphological variation within the species and to aggressive breeding practices that have resulted in inbred populations of dogs. Many breeds of dog are particularly susceptible to inherited diseases that are also common in humans, such as cancer, heart disease, rheumatoid arthritis, autoimmune disorders, deafness, and blindness. The size of the haploid dog genome is estimated to be 2445 Mb. The diploid genome is organized in 38 pairs of autosomes and two sex chromosomes.

Project type: Organism overview

Lineage: *Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae; Canis; Canis lupus; Canis lupus familiaris*

Component Projects

Project Type	Number of Projects
<ul style="list-style-type: none"> ▶ Genome sequencing 1 Chromosome(s); 1 Contigs; 1 Unknown 	3
<ul style="list-style-type: none"> ▶ Map 	2
<ul style="list-style-type: none"> ▶ RefSeq Genome 1 Chromosome(s) 	1

Project Data

PMC: 1
 Pubmed: 6
 Genome: 39
 Nucleotide: 4356






NCBI Links

- [MapViewer](#)
- [BLAST](#)
- [Dog Genome Resources](#)

Related Resources

- [Canine Genetic Linkage Map](#)
- [The FHCRC Dog Genome Project](#)
- [The Dog Genome Project, UC Berkeley](#)
- [The Canine Radiation Hybrid Project](#)
- [Cytogenetics and Canine Cancer, NC State](#)
- [Canine Sequencing White Paper](#)
- [Online Mendelian Inheritance in Animals](#)
- [DogMap](#)
- [NISC Comparative Vertebrate Sequencing](#)

Recent activity

- [Turn Off](#) [Clear](#)
-  [NFATC1 \[Canis lupus familiaris\]](#) Gene
 -  [Canis lupus familiaris](#) BioProject (Genome Project)
 -  [Opn1mw opsin 1 \(cone pigments\), medium-wave-sensitive \(color blindness, deutan\) ...](#) Gene
 -  [opsin \(1340\)](#) Gene
 -  [txid9612\[orgn\] \(31566\)](#) Gene

[See more](#)

Gene

Gene

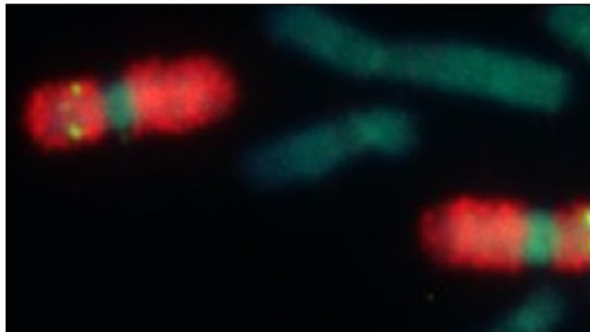
human opsin



Search

[Limits](#) [Advanced](#)

[Help](#)



Welcome to Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

Using Gene

[Gene Quick Start](#)

[FAQ](#)

[Download/FTP](#)

[RefSeq Mailing List](#)

[Gene News](#) 

Gene Tools

[Submit GeneRIFs](#)

[Submit Correction](#)

[Statistics](#)

[BLAST](#)

[Genome Workbench](#)

[Splign](#)

Other Resources

[HomoloGene](#)

[OMIM](#)

[RefSeq](#)

[RefSeqGene](#)

[UniGene](#)

[Protein Clusters](#)

NCBI Entrez Gene: Human Opsin

<http://www.ncbi.nlm.nih.gov/gene?term=human%20opsin>

Gene [Save search](#) [Limits](#) [Advanced](#) [Help](#)

[Display Settings:](#) Summary, 20 per page, Sorted by Relevance [Send to:](#)

Results: 1 to 20 of 58

<< First < Prev Page of 3 [Next >](#) [Last >>](#)

[OPN1LW](#)

1. **Official Symbol:** OPN1LW and **Name:** opsin 1 (cone pigments), long-wave-sensitive [*Homo sapiens*]
Other Aliases: hCG_41347, CBBM, CBP, COD5, RCP, ROP
Other Designations: OTTHUMP00000032193; cone dystrophy 5 (X-linked); long-wave-s
Chromosome: X; **Location:** Xq28
Annotation: Chromosome X, NC_000023.10 (153409725..153424507)
MIM: 300822
 ID: 5956
[Order cDNA clone](#)

[OPN1SW](#)

2. **Official Symbol:** OPN1SW and **Name:** opsin 1 (cone pigments), short-wave-sensitive [*Homo sapiens*]
Other Aliases: BCP, BOP, CBT
Other Designations: OTTHUMP00000212782; blue cone photoreceptor pigment; blue-s
Chromosome: 7; **Location:** 7q32.1
Annotation: Chromosome 7, NC_000007.13 (128412543..128415844, complement)
MIM: 613522
 ID: 611
[Order cDNA clone](#)

Filter your results:

All (58)

[Current Only \(52\)](#)

[Genes Genomes \(50\)](#)

[SNP GeneView \(44\)](#)

[In Variation Viewer \(8\)](#)

[Manage Filters](#)

▼ **Top Organisms [Tree]**

[Homo sapiens \(24\)](#)

[Mus musculus \(20\)](#)

[Rattus norvegicus \(6\)](#)

[Pediculus humanus corporis \(3\)](#)

[Bos taurus \(2\)](#)

[All other taxa \(3\)](#)

[More...](#)

Find related data

Database:

Entrez Gene: Human Opsin OPN1MW

<http://www.ncbi.nlm.nih.gov/gene/2652>

Gene

Gene

Search

Limits Advanced

Help

Display Settings: Full Report

Send to:

OPN1MW opsin 1 (cone pigments), medium-wave-sensitive [*Homo sapiens*]

Gene ID: 2652, updated on 29-Sep-2011

Summary

Official Symbol OPN1MW provided by [HGNC](#)

Official Full Name opsin 1 (cone pigments), medium-wave-sensitive provided by [HGNC](#)

Primary source [HGNC:4206](#)

See related [Ensembl:ENSG00000147380](#); [HPRD:02365](#); [MIM:300821](#)

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as CBD; GCP; GOP; CBBM; COD5; OPN1MW1; OPN1MW2; MGC176615; MGC177321; MGC198468; MGC198469

Summary This gene encodes for a light absorbing visual pigment of the opsin gene family. The encoded protein is called green cone photopigment or medium-wavelength sensitive opsin. Opsins are G-protein coupled receptors with seven transmembrane domains, an N-terminal extracellular domain, and a C-terminal cytoplasmic domain. The long-wavelength opsin gene and multiple copies of the medium-wavelength opsin gene are tandemly arrayed on the X chromosome and frequent unequal recombination and gene conversion may occur between these sequences. X chromosomes may have fusions of the medium- and long-wavelength opsin genes or may have more than one copy of these genes. Defects in this gene are the cause of deutanopic colorblindness. [provided by RefSeq, Mar 2009]

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- Phenotypes
- General gene info
- General protein info
- Reference sequences
- Related sequences
- Additional links

Links

- Order cDNA clone
- BioAssay, by Gene target
- BioAssays, Gene target, Active
- BioProjects
- BioSystems
- Books
- CCDS
- Conserved Domains
- dbVar
- Full text in PMC
- Genome

Entrez Gene: Human Opsin OPN1MW

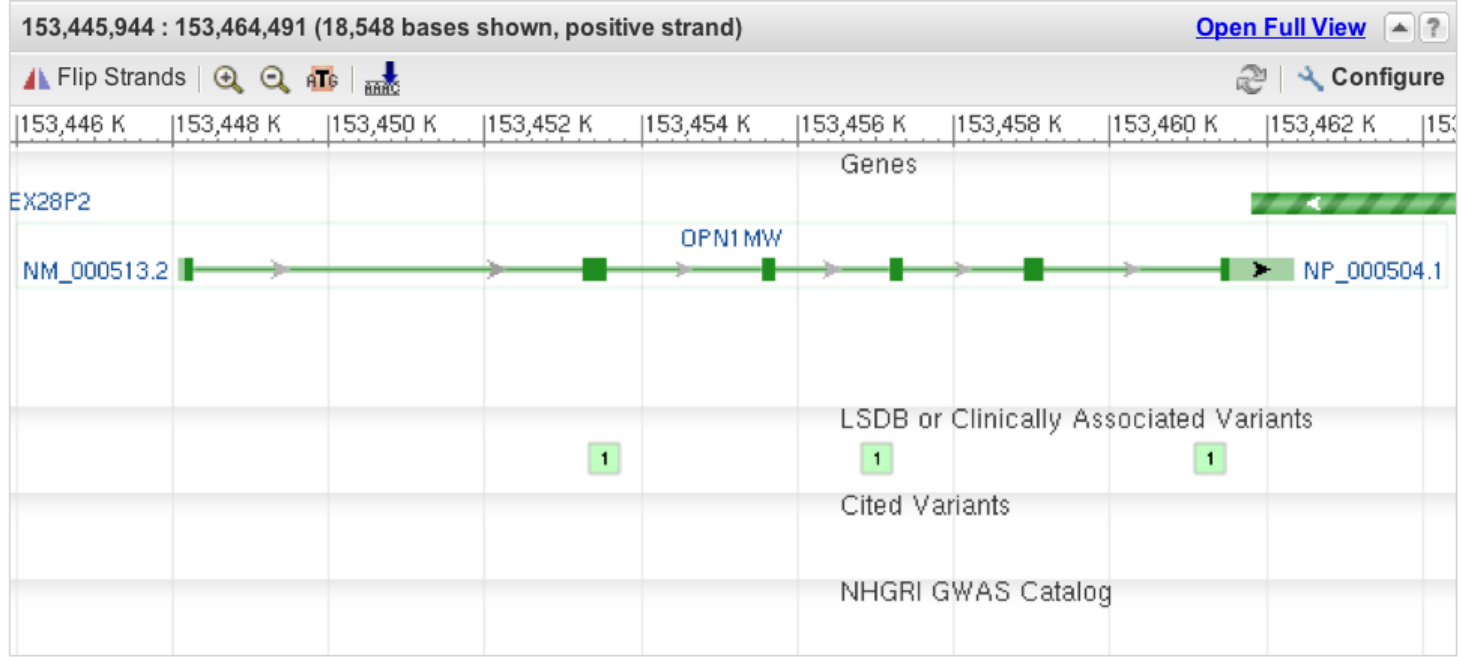
<http://www.ncbi.nlm.nih.gov/gene/2652>

Genomic regions, transcripts, and products

Go to [reference sequence details](#)

Genomic Sequence

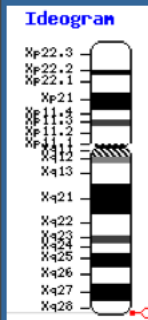
Go to nucleotide [Graphics](#) [FASTA](#) [GenBank](#)



- [PubChem Substance](#)
- [PubMed](#)
- [PubMed \(GeneRIF\)](#)
- [PubMed \(OMIM\)](#)
- [RefSeq Proteins](#)
- [RefSeq RNAs](#)
- [RefSeqGene](#)
- [SNP](#)
- [SNP: GeneView](#)
- [SNP: Genotype](#)
- [SNP: VarView](#)
- [Taxonomy](#)
- [UniGene](#)
- [UniSTS](#)

Links to other resources

- [AceView](#)
- [Ensembl](#)
- [Evidence Viewer](#)
- [GeneTests for MIM: 300821](#)
- [GeneTests for MIM: 303800](#)
- [HGNC](#)



Homo sapiens (human) Build 37.3 (Current)

[BLAST human sequences](#)

Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 [X] Y MT

Query: 2652[*gene_id*] [clear]

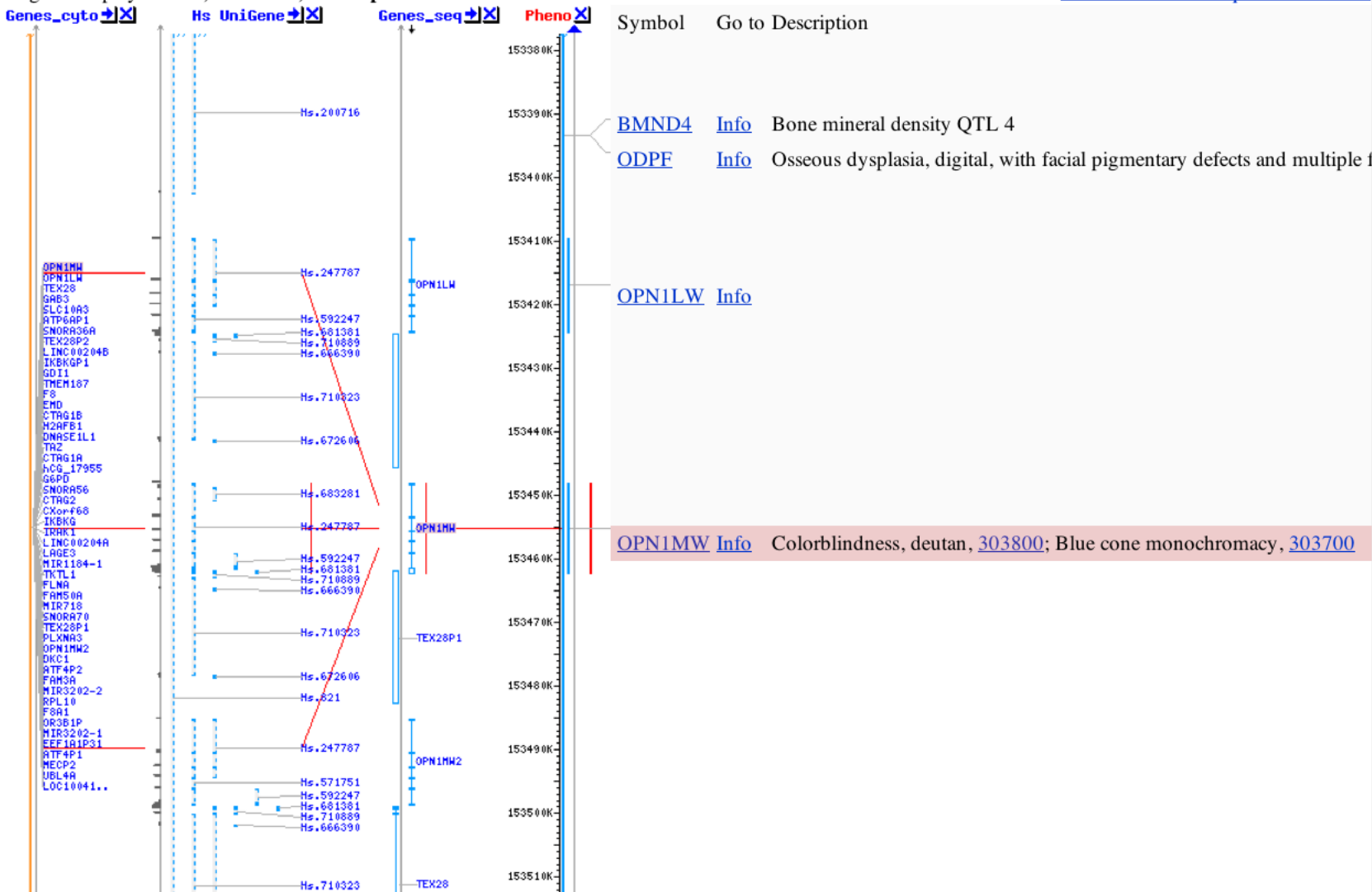
Master Map: Phenotype (includes QTLs)

[Summary of Maps](#)

[Maps & Options](#)

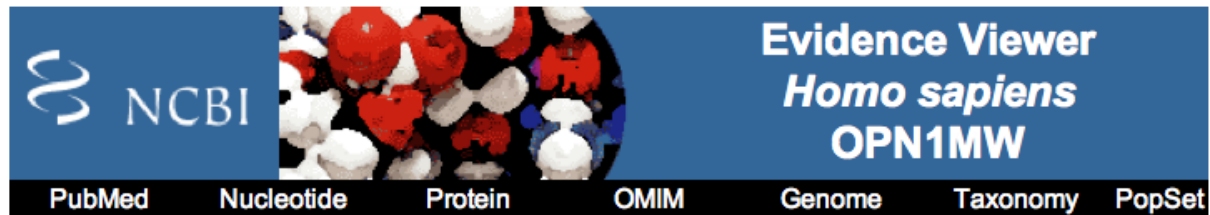
Region Displayed: 153,378K-153,533K bp

[Download/View Sequence/Evidence](#)



Evidence Viewer for OPN1MW

http://www.ncbi.nlm.nih.gov/sutils/evv.cgi?taxid=9606&contig=NT_167198.1&gene=OPN1MW



NCBI Evidence Viewer
Homo sapiens
OPN1MW

PubMed Nucleotide Protein OMIM Genome Taxonomy PopSet

Key for display of mRNAs aligning in this region:

[MapView](#)

[Evidence Viewer Help](#)

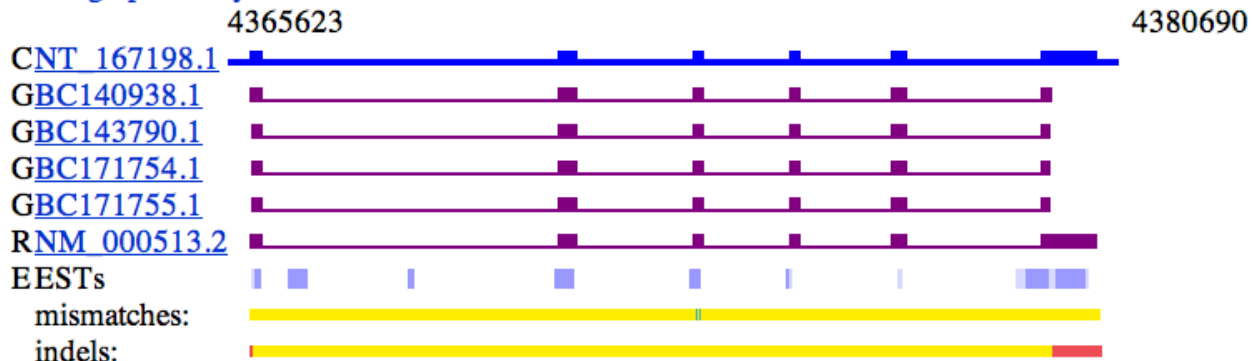
- Genomic sequence (C)
 - model exons, single (M)
 - mRNA exons, single (G, R)
 - model exons, [overlapping](#) (M)
 - mRNA exons, [overlapping](#) (G, R)
- C = contig; M = model mRNA; R = RefSeq mRNA; G = GenBank mRNA
R = new since last genome build; **R** = updated since last genome build

EST density key (E):

- 1 EST
- 2-5 ESTs
- 6-20 ESTs
- 21-99 ESTs
- >100 ESTs

6 exons and 1 gene found in this genomic region spanning 15068 bp.


[View graphic only](#)



Mouse over mismatches, indels and unaligned regions to see their exon number.

OMIM Home Page

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>



The screenshot shows the OMIM Home Page interface. At the top, there is a search bar with the text "Search OMIM for" and buttons for "Go" and "Clear". Below the search bar are navigation options: "Limits", "Preview/Index", "History", "Clipboard", and "Details". The page is divided into a left sidebar and a main content area. The sidebar contains links for "Entrez", "OMIM", "Help", "FAQ", "OMIM Facts", "Allied Resources", and "Human Genome Resources". The main content area features a list of search instructions and a section titled "OMIM™ - Online Mendelian Inheritance in Man™" with a "NEW" announcement.

Search OMIM for

[Limits](#) [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)

Entrez

- OMIM
 - Search OMIM
 - Search Gene Map
 - Search Morbid Map
- Help
 - OMIM Help
 - How to Link
- FAQ
 - Numbering System
 - Symbols
 - How to Print
 - Citing OMIM
 - Download
- OMIM Facts
 - Statistics
 - Update Log
 - Restrictions on Use
- Allied Resources
 - Genetic Alliance
 - Databases
 - HGMD
 - Locus-Specific
 - Model Organisms
 - MitoMap
 - Phenotype
 - Davis Human/Mouse
 - Homology Maps
 - Coriell
 - The Jackson Laboratory
 - Human Gene
 - Nomenclature
- Human Genome Resources
 - Genes and Disease
 - LocusLink
 - Map Viewer
 - Sequencing Progress

- Enter one or more search terms.
- Use **Limits** to restrict your search by search field, chromosome, and other criteria.
- Use **Index** to browse terms found in OMIM records.
- Use **History** to retrieve records from previous searches, or to combine searches.

OMIM™ - Online Mendelian Inheritance in Man™

NEW OMIM is now incorporated into NCBI's Entrez system and can be queried using the same approach as the other Entrez databases such as PubMed and GenBank. The previous OMIM pages are still available [here](#).

Welcome to OMIM, Online Mendelian Inheritance in Man. This database is a catalog of human genes and genetic disorders authored and edited by Dr. Victor A. McKusick and his colleagues at Johns Hopkins and elsewhere, and developed for the World Wide Web by NCBI, the National Center for Biotechnology Information. The database contains textual information and references. It also contains copious links to MEDLINE and sequence records in the Entrez system, and links to additional related resources at NCBI and elsewhere.

You can do a search by entering one or more terms in the text box above. Advanced search options are accessible via the Limits, Preview/Index, History, and Clipboard options in the grey bar beneath the text box. The [OMIM help](#) document provides additional information and examples of basic and advanced searches.

The links to the left provide further technical information, searching options, frequently asked questions ([FAQ](#)), and information on allied resources. To return to this page, click on the OMIM link in the black header bar or on the graphic at the top of any OMIM page.

NOTE: OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are urged to consult with a qualified physician for diagnosis and for answers to personal questions.

Colorblindness in OMIM

All Databases PubMed Nucleotide Protein Genome Structure PMC OMIM

Search for

Display Show






All: 3 

Items 1 - 3 of 3

One page.

- 1: [#303800. COLORBLINDNESS, PARTIAL, DEUTAN SERIES; CBD](#)** GeneTests, Links
 DEUTERANOMALY, INCLUDED
 Gene map locus [Xq28](#)
- 2: [#303700. BLUE CONE MONOCHROMACY; BCM](#)** GeneTests, Links
 CONE DYSTROPHY 5, X-LINKED, INCLUDED; COD5, INCLUDED
 Gene map locus [Xq28](#), [Xq28](#)
- 3: [*300821. OPSIN 1, MEDIUM-WAVE-SENSITIVE; OPN1MW](#)** MGI, GeneTests, Links
 Gene map locus [Xq28](#)

Recent activity 

- [Turn Off](#) [Clear](#)
-  [OMIM for Gene \(Select 265... \(3\) OMIM](#)
 -  HomoloGene:88332.Gene exclusive to H.sapiens
 -  [HomoloGene for Gene \(Sele... \(1\) HomoloGene](#)
 -  [BioSystems for Gene \(Sele... BioSystems](#)
 -  Acanthocheilonema viteae
- » See more...

Colorblindness in OMIM

<http://omim.org/entry/303800>

#303800

COLORBLINDNESS, PARTIAL, DEUTAN SERIES; CBD

Alternative titles; symbols

DEUTAN COLORBLINDNESS; DCB

DEUTERANOPIA

GREEN COLORBLINDNESS

Other entities represented in this entry:

DEUTERANOMALY, INCLUDED

Phenotype Gene Relationships

Location	Phenotype	Phenotype MIM number	Gene/Locus	Gene/Locus MIM number
Xq28	Colorblindness, deutan	303800	OPN1MW	300821

Clinical Synopsis

TEXT

A number sign (#) is used with this entry because deutan colorblindness is caused by mutation in the OPN1MW gene ([300821](#)), which encodes green cone pigment.

Description

Normal color vision in humans is trichromatic, being based on 3 classes of cone that are maximally sensitive to light at approximately 420 nm (blue cones; [613522](#)), 530 nm (green cones; [300821](#)), and 560 nm (red cones; [300822](#)). Comparison by neural circuits of light absorption by the 3 classes of cone photoreceptors allows perception of red, yellow, green, and blue colors individually or in various combinations. Dichromatic color vision is severely defective color vision based on the use of only 2 types of photoreceptors, blue plus green (protanopia; see [303900](#)) or blue plus red (deuteranopia). Anomalous trichromacy is trichromatic color vision based on a blue, green, and an anomalous red-like photoreceptor (protanomaly), or a blue, red, and an anomalous green-like photoreceptor (deuteranomaly). The color vision defect is generally mild but may in

[Table of Contents - #303800](#)

External Links:

▼ [Clinical Resources](#)

[Clinical Trials](#)

[Gene Tests](#)

[Genetic Alliance](#)

[Genetics Home Reference](#)


► [Variation](#)

► [Animal Models](#)

► [Cellular Pathways](#)

Human Genome Resources

<http://www.ncbi.nlm.nih.gov/genome/guide/human/>


 NCBI • Genomic Biology • Homo sapiens
 Search for


Browse your Genome
Click on the Chromosome to show


Genes



Find A Gene
Search for
from

The NCBI Handbook

 An online guide to the use of NCBI resources. Titles of selected chapters that refer to human genome resources are shown below.

 **The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation**
Adrienne Kitts and Stephen Sherry



A challenge facing researchers today is that of piecing together and analyzing the plethora of data currently being generated through the Human Genome Project and scores of smaller projects. NCBI's Web site serves an integrated, one-stop, genomic information infrastructure for biomedical researchers from around the world so that they may use these data in their research efforts. [More...](#)

Genes and Human Health

▶ Gene Database

A new database of genes and associated information is now available for searching in Entrez.

▶ dbSNP

A database of single nucleotide polymorphisms (SNPs) and other nucleotide variations.

▶ OMIM

A guide to human genes and inherited disorders maintained by Johns Hopkins University and collaborators.

▶ dbGaP

The database of Genotypes and Phenotypes (dbGaP) was developed to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype.

Epigenomics

▶ NIH Epigenomics Roadmap

Reference epigenomic maps and studies on new epigenetic mechanisms and their relevance to human health.

▶ Roadmap Epigenomics Data

A comprehensive listing of all NIH Roadmap Epigenomics datasets submitted to GEO and SRA.



[PubMed](#)
[All Databases](#)
[BLAST](#)
[OMIM](#)
[Books](#)
[Taxonomy](#)
[Structure](#)

Search for

- [Brief Description](#)
- [Scope](#)
- [Announcements](#)
- [Access and Availability](#)
- [Distinguishing Features](#)
- [References](#)

NCBI Reference Sequences

The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins. RefSeq is a foundation for medical, functional, and diversity studies; they provide a stable reference for genome annotation, gene identification and characterization, mutation and polymorphism analysis (especially [RefSeqGene](#) records), expression studies, and comparative analyses. [\[more...\]](#)

▸ Scope ↑


NCBI provides RefSeqs for taxonomically diverse organisms including eukaryotes, bacteria, and viruses. Additional records are added to the collection as data become publicly available.

▸ Announcements ↑

September 7, 2011: The Conserved CDS (CCDS) database released an update for human following genome reannotation by NCBI and Ensembl. This update adds 972 new CCDS IDs, bringing the total to 26,473 consistently consistently annotated coding regions that pass CCDS QA tests. [\[more\]](#)

Site contents

Information

[NCBI Handbook](#)
[Overview](#) | [FAQ](#) 
[Accessions](#) | [Status](#) |
[Queries](#) | [Publications](#)

FTP

[RefSeq Release](#)
[Catalog](#) | [Notes](#)
[Genomes](#)
[BLAST databases](#)

Statistics

[Release Statistics](#)

Feedback

[NCBI Help Desk](#)
[Submit Updates](#)
[Submit GeneRIF](#)

Subscribe to email list or RSS

[RefSeq](#) | [Gene News \(RSS\)](#)
[Map Viewer](#) | [NCBI](#)

Related links

[Genomic Biology Home](#)
[Gene](#) | [Genome Project](#)
[Entrez Genomes Home](#)
[Map Viewer](#) | [UniGene](#)

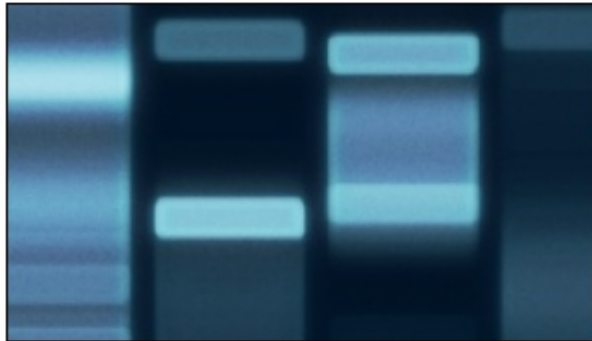
Credits

[Collaborators](#)
[Microbial Providers](#)
[Viral Genome Advisors](#)

RefSeqGene

Gene

Search



RefSeqGene

RefSeqGene defines genomic sequences to be used as reference standards for well-characterized genes and is part of the Locus Reference Genomic (LRG) Project.

Using RefSeqGene

- [About](#)
- [Browse Genes with RefSeqGene Sequences](#)
- [See RefSeqGene Sequences in the Nucleotide Database](#)
- [Download Data](#)
- [New RefSeqGene Sequences \(last 30 days\)](#)

Tools

- [Clinical Remap](#)
- [Genome Workbench](#)
- [Map Viewer](#)
- [RefSeqGene BLAST](#)

Related Sites

- [Entrez Gene](#)
- [RefSeq](#)
- [Locus Reference Genomic \(LRG\)](#)
- [Variation Databases \(dbSNP, dbVar\)](#)
- [Online Mendelian Inheritance In Man](#)
- [Contact Us](#)



Search UniGene for

[Limits](#) [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)

UniGene
[Homepage](#)
[FAQs](#)
[Query Tips](#)
[Library Browser](#)
[DDD](#)
[Download UniGene](#)

Related Databases
[Gene](#)
[HomoloGene](#)
[dbEST](#)
[Trace Archive](#)

NIH cDNA Projects
[MGC](#) | [ZGC](#) | [XGC](#)
[Finding cDNAs](#)

UniGene: An Organized View of the Transcriptome.

Each UniGene entry is a set of transcript sequences that appear to come from the same transcription locus (gene or expressed pseudogene), together with information on protein similarities, gene expression, cDNA clone reagents, and genomic location.

Species UniGene Entries

Chordata

Mammalia



Bos taurus (cow)	43,448
Canis lupus familiaris (dog)	27,853
Equus caballus (horse)	8,348
Homo sapiens (human)	123,891
Macaca fascicularis (crab-eating macaque)	11,980
Macaca mulatta (rhesus monkey)	15,359
Monodelphis domestica (gray short-tailed opossum)	359
Mus musculus (mouse)	79,119
Ornithorhynchus anatinus (platypus)	1,831
Oryctolagus cuniculus (rabbit)	6,576
Ovis aries (sheep)	14,659
Papio anubis (olive baboon)	12,110
Pongo abelii (Sumatran orangutan)	6,996
Rattus norvegicus (Norway rat)	65,511
Sus scrofa (pig)	51,670
Trichosurus vulpecula (silver-gray brushtail possum)	11,771

Actinopterygii

Danio rerio (zebrafish)	56,944
Fundulus heteroclitus (killifish)	4,618
Gadus morhua (Atlantic cod)	14,542
Gasterosteus aculeatus (three spined stickleback)	18,938
Oncorhynchus mykiss (rainbow trout)	25,025
Oryzias latipes (Japanese medaka)	22,552
Pimephales promelas (fathead minnow)	21,765
Salmo salar (Atlantic salmon)	36,314
Takifugu rubripes (pufferfish)	3,809

Amphibia

Xenopus laevis (African clawed frog)	35,077
--	--------

[All Databases](#)
[PubMed](#)
[Nucleotide](#)
[Protein](#)
[Genome](#)
[Structure](#)
[OMIM](#)
[PMC](#)
[Journals](#)
[Books](#)

Search for

[Limits](#)
[Preview/Index](#)
[History](#)
[Clipboard](#)
[Details](#)

HomoloGene

[Homepage](#)
[Query Tips](#)
[Build Procedure](#)
[FTP site](#)

Genome

Resources

[Homo sapiens](#)
[Mus musculus](#)
[Rattus norvegicus](#)
[Danio rerio](#)

HomoloGene is a system for automated detection of homologs among the annotated genes of several completely sequenced eukaryotic genomes.

HomoloGene Release 65 Statistics

Initial numbers of genes from complete genomes, numbers of genes placed in a homology group, and the numbers of groups for each species.

Species	Number of Genes		HomoloGene groups
	Input	Grouped	
Homo sapiens	19,943*	18,981	18,431
Pan troglodytes	25,096	16,850	15,980
Canis familiaris	19,766	16,708	15,951
Bos taurus	22,049	18,180	16,224
Mus musculus	25,388	21,766	19,005
Rattus norvegicus	21,991	19,229	17,473
Gallus gallus	17,959	13,142	11,905
Danio rerio	26,690*	21,084	14,067
Drosophila melanogaster	13,827*	9,282	7,749
Anopheles gambiae	12,460	8,867	7,541
Caenorhabditis elegans	20,132*	8,678	4,810
Schizosaccharomyces pombe	5,043	3,225	2,935
Saccharomyces cerevisiae	5,880	4,851	4,370
Kluyveromyces lactis	5,335	4,459	4,382
Eremothecium gossypii	4,722	3,928	3,884
Magnaporthe grisea	12,832	7,330	6,399
Neurospora crassa	9,821*	6,287	6,144
Arabidopsis thaliana	27,309*	19,961	11,243
Oryza sativa	26,887	17,276	10,627
Plasmodium falciparum	5,266	1,862	799

* indicates organisms where new genome annotation data is used in this build.

Last updated on: Mon Feb 14 2011

What's New

HomoloGene release 65 includes updated annotations for the following species: Homo sapiens (NCBI release 37.2), Danio rerio (NCBI release 4.1), Drosophila melanogaster (NCBI release 9.3) Caenorhabditis elegans (NCBI release 9.1), Arabidopsis thaliana (NCBI release 9.1).

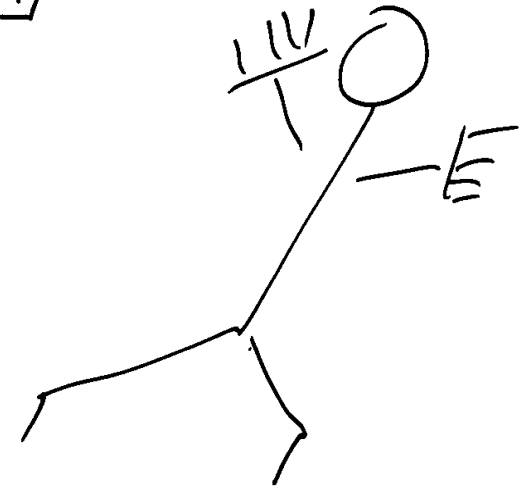
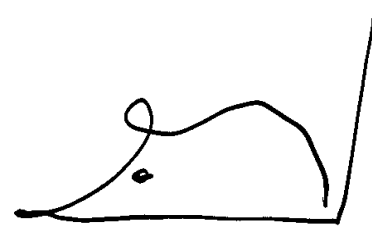
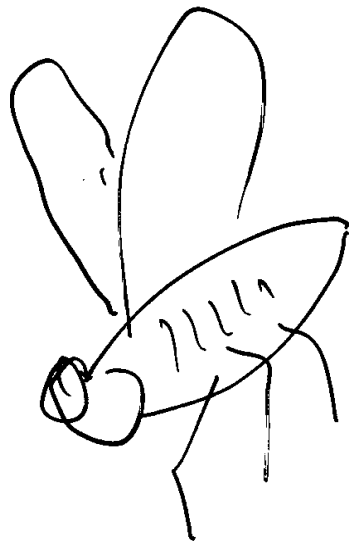
Related Resources

Entrez Genomes

A collection of complete genome sequences that includes more than 1000 viruses and over hundred microbes

- [Archaea](#)
- [Bacteria](#)
- [Eukaryota](#)
- [Viruses](#)

Comparative Genomics



L'homme et ses Cousins germains
F. Jacob

Ensembl Home Page

<http://www.ensembl.org/>

Search: for
 e.g. [BRCA2](#) or [rat X:100000..200000](#) or [coronary heart disease](#)

Browse a Genome

The Ensembl project produces genome databases for vertebrates and other eukaryotic species, and makes this information freely available online.

Click on a link below to go to the species' home page.

Favourite genomes [\(Change favourites\)](#)



Human
GRCh37



Mouse
NCBIM37



Zebrafish
Zv9

All genomes

-- Select a species --

[View full list of all Ensembl species](#)

Other species are available in [Ensembl Pre!](#) and [EnsemblGenomes](#)



Ensembl is a joint project between [EMBL - EBI](#) and the [Wellcome Trust Sanger Institute](#) to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes.



Ensembl receives major funding from the Wellcome Trust. Our [acknowledgements page](#) includes a list of additional current and previous funding bodies.

New to Ensembl?

Did you know you can:

- [Learn how to use Ensembl](#)
with our video tutorials and walk-throughs
- [Add custom tracks](#)
using our new Control Panel
- [Upload and analyse your data](#)
and save it to your Ensembl account
- [Search for a DNA or protein sequence](#)
using BLAST or BLAT
- [Fetch only the data you want](#)
from our public database, using the Perl API
- [Download our databases via FTP](#)
in FASTA, MySQL and other formats
- [Mine Ensembl with BioMart](#)
and export sequences or tables in text, html, or Excel format

Still got questions? Try our [FAQs](#) or [glossary](#)

Did you know...?

CpG Find CpG islands in [region in detail](#). Turn them on using [configure this page](#).

What's New in Release 64 (September 2011)

- [Human assembly updated to GRCh37.5](#)
- [New regulatory data \(Human, Mouse\)](#)
- [Auto-resizing images \(all species\)](#)

[Full details of this release](#)

[More release news on our blog →](#)

Latest blog posts

- [Human Methylation data in Ensembl](#)
- [Ensembl Genomes Release 11](#)
- [Workshop on automatic gene annotation](#)

[Go to Ensembl blog →](#)

EBI Genomes Home Page

<http://www.ensembl.org/>

Your Ensembl

- [Login or Register](#)
- [About User Accounts](#)

Help & Documentation

- [Table of Contents](#)
- [Helpdesk](#)
- [What's New](#)
- [About Ensembl](#)
- [Downloading data](#)
- [Displaying your own data](#)
- [Ensembl software](#)

Select a species

- [Mammals](#)
- [Other chordates](#)
- [Other eukaryotes](#)

Ensembl Archive

- [View previous release of page in Archive!](#)
- [Stable Archive! link for this page](#)



Search Ensembl

Search: for

e.g. mouse chromosome 2 or X:10000..20000 or human gene BRCA2

Ensembl tools



Start a sequence search →
Search Ensembl for nucleotide and peptide sequences with BLAST and SSAHA.



Mine Ensembl with BioMart →
Cross-reference Ensembl datasets with BioMart, a powerful data-mining tool.



Customise Your Ensembl →
Register with Ensembl to bookmark your favourite pages, customise your home page and much more!



Fetch data with the Ensembl API →
Learn how to extract data from the public Ensembl database with this tutorial.

Ensembl 42

Pre! species

Popular genomes



Homo sapiens
NCBI 36 | Vega



Mus musculus
NCBI m36 | Vega



Danio rerio
Zv6 | Vega

More genomes

- ▶ **Aedes aegypti** AaegL1
- ▶ **Anopheles gambiae** AgamP3
- ▶ **Bos taurus** Btau 2.0
- ▶ **Caenorhabditis elegans** WS160
- ▶ **Canis familiaris** CanFam 2.0 **UPDATED!**
- ▶ **Ciona intestinalis** JGI 2
- ▶ **Ciona savignyi** CSAV 2.0
- ▶ **Dasytus novemcinctus** ARMA
- ▶ **Drosophila melanogaster** BDGP 4.3

Ensembl headlines: Release 42 (December 2006)

[New - User accounts](#) (all species)

Ensembl Human Genome


http://www.ensembl.org/Homo_sapiens/

Human (GRCh37) ▾


About this species

Description

- [-] Genome Statistics
 - Assembly and Genebuild
 - Top 40 InterPro hits
 - Top 500 InterPro hits
- What's New
- [-] Sample entry points
 - Karyotype
 - Location (6:133017695-1331
 - Gene (BRCA2)
 - Transcript (FOXP2-203)
 - Variation (rs1333049)
 - Regulation (ENSR00001348

 Configure this page

 Manage your data

 Export data

 Bookmark this page

Search Ensembl Human

Search for:
 e.g. [BRCA2](#) or [6:133017695-133161157](#) or [osteoarthritis](#)

Description

Human (*Homo sapiens*)

Assembly

This site provides a data set based on the February 2009 *Homo sapiens* high coverage assembly GRCh37 (GCA_000001405.6) from the [Genome Reference Consortium](#). This assembly is used by UCSC to create their hg19 database. The data set consists of gene models built from the genewise alignments of the human proteome as well as from alignments of human cDNAs using the cDNA2genome model of exonerate.

This release of the assembly has the following properties:

- 27478 contigs.
- contig length total 3.2 Gb.
- chromosome length total 3.1 Gb.

It also includes nine [haplotypic regions](#), mainly in the MHC region of chromosome 6.

As the GRC maintains and improves the assembly, patches are being introduced. [Patch release five \(GRCh37.p5\)](#) was included in Ensembl release 64. Currently, assembly patches are of two types:

- Novel patch: new sequences that add alternative sequence at a loci and will remain as haplotypes in the next major assembly release by GRC
- Fix patch: sequences that correct the reference sequence and will replace the given region of the reference assembly at the next major assembly release by GRC



Ensembl Human Opsin Search

http://uswest.ensembl.org/Homo_sapiens/Search/Results?species=Homo_sapiens;idx=;q

Human (GRCh37) ▾

Search Ensembl

[New Search](#)

[Configure this page](#)

[Manage your data](#)

[Export data](#)

[Bookmark this page](#)

Results Summary

Your search of Human with 'opsin' returned the following results:

By Feature type

Total	56
▶ Domain	8
▶ Family	6
▶ Gene	9
▶ Transcript	33

By Species

Total	56
▶ Human	56

Ensembl release 64 - Sep 2011 © [WTSI](#) / [EBI](#)

[About Ensembl](#) | [Contact Us](#) | [Help](#)

[Permanent link](#) - [View in archive site](#)

Ensembl Human Opsin Genes

http://uswest.ensembl.org/Homo_sapiens/Search/Results?species=Homo_sapiens;idx=;q=



Human (GRCh37) ▾

Search Ensembl

[New Search](#)

[Configure this page](#)

[Manage your data](#)

[Export data](#)

[Bookmark this page](#)

Result in Detail

9 Genes match your query ('opsin') in Human

OPN1LW [Ensembl/Havana merge: ENSG00000102076]

Description opsin 1 (cone pigments), long-wave-sensitive [Source:HGNC Symbol;Acc:9936] [Type: protein coding Ensembl/Havana merge]
Location [X:153409698-153424507:1](#)
Source e64

OPN1SW [Ensembl/Havana merge: ENSG00000128617]

Description opsin 1 (cone pigments), short-wave-sensitive [Source:HGNC Symbol;Acc:1012] [Type: protein coding Ensembl/Havana merge]
Location [7:128412545-128415844:-1](#)
Source e64

OPN5 [Ensembl/Havana merge: ENSG00000124818]

Description opsin 5 [Source:HGNC Symbol;Acc:19992] [Type: protein coding Ensembl/Havana merge]
Location [6:47749718-47800516:1](#)
Source e64

OPN1MW [Ensembl/Havana merge: ENSG00000147380]

Description opsin 1 (cone pigments), medium-wave-sensitive [Source:HGNC Symbol;Acc:4206] [Type: protein coding Ensembl/Havana merge]
Location [X:153448107-153461633:1](#)
Source e64



Human (GRCh37) ▾

Location: X:153,448,107-153,461,633

Gene: OPN1MW

Gene-based displays

- Gene summary
- Splice variants (3)
- Supporting evidence
- Sequence
- External references
- Regulation
- Comparative Genomics
 - Genomic alignments
 - Gene Tree (image)
 - Gene Tree (text)
 - Gene Tree (alignment)
 - Orthologues (82)
 - Paralogues (9)
 - Protein families (1)
- Phenotype
- Genetic Variation
 - Variation Table
 - Variation Image
 - Structural Variation
- External Data
 - Personal annotation
- ID History
 - Gene history

- Configure this page
- Manage your data
- Export data
- Bookmark this page

Gene: OPN1MW (ENSG00000147380)

Description opsin 1 (cone pigments), medium-wave-sensitive [Source:HGNC Symbol;Acc:4206]

Location [Chromosome X: 153,448,107-153,461,633](#) forward strand.

Transcripts ☐ There are 3 transcripts in this gene

Show/hide columns		Filter				
Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
OPN1MW-001	ENST00000369935	1257	ENSP00000358951	364	Protein coding	CCDS14743
OPN1MW-002	ENST00000430054	590	ENSP00000394838	164	Protein coding	-
OPN1MW-003	ENST00000468495	692	No protein product	-	Processed transcript	-

Transcript and Gene level displays [X]

In Ensembl we provide displays at two levels:

- Transcript views which provide information specific to an individual transcript such as the cDNA and CDS sequences and protein domain annotation.
- Gene views which provide displays for data associated at the gene level such as orthologues, paralogues, regulatory regions and splice variants.

This view is a gene level view. To access the transcript level displays select a Transcript ID in the table above and then navigate to the information you want using the menu at the left hand side of the page. To return to viewing gene level information click on the Gene tab in the menu bar at the top of the page.

Gene summary [help](#)

Name [OPN1MW](#) (HGNC Symbol)

Synonyms CBBM, CBD, COD5, GCP, OPN1MW1 [To view all Ensembl genes linked to the name [click here.](#)]

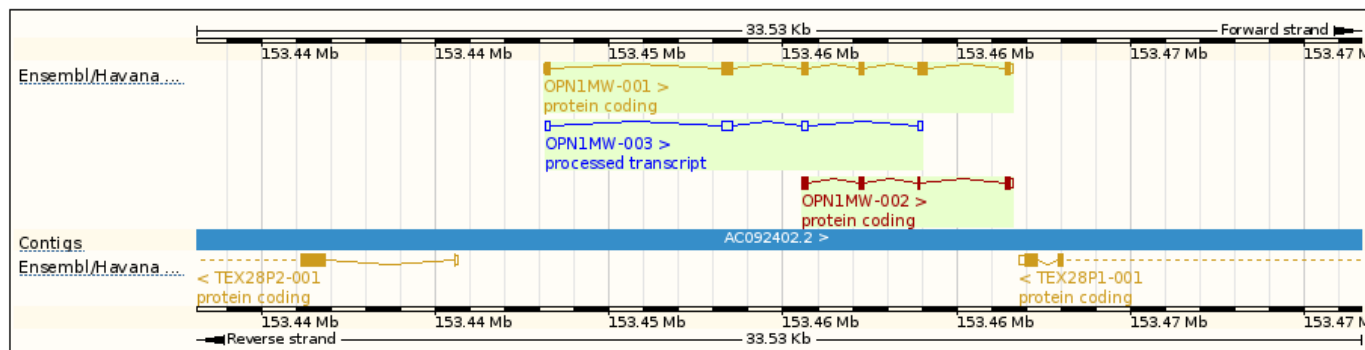
CCDS This gene is a member of the Human CCDS set: [CCDS14743](#)

Gene type Known protein coding

Prediction Method Annotation for this gene includes both automatic annotation from Ensembl and [Havana](#) manual curation, see [article](#).

Alternative genes This gene corresponds to the following database identifiers:

Havana gene: [OTTHUMG00000022652](#) (version 6) [[view all locations](#)]

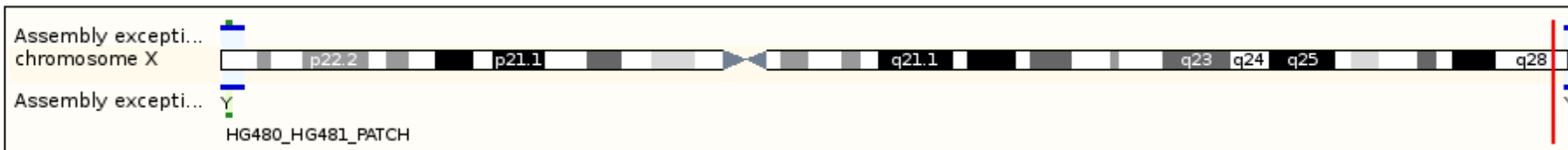


[Export Image](#)

Ensembl Opsin OPN1MW Gene Location

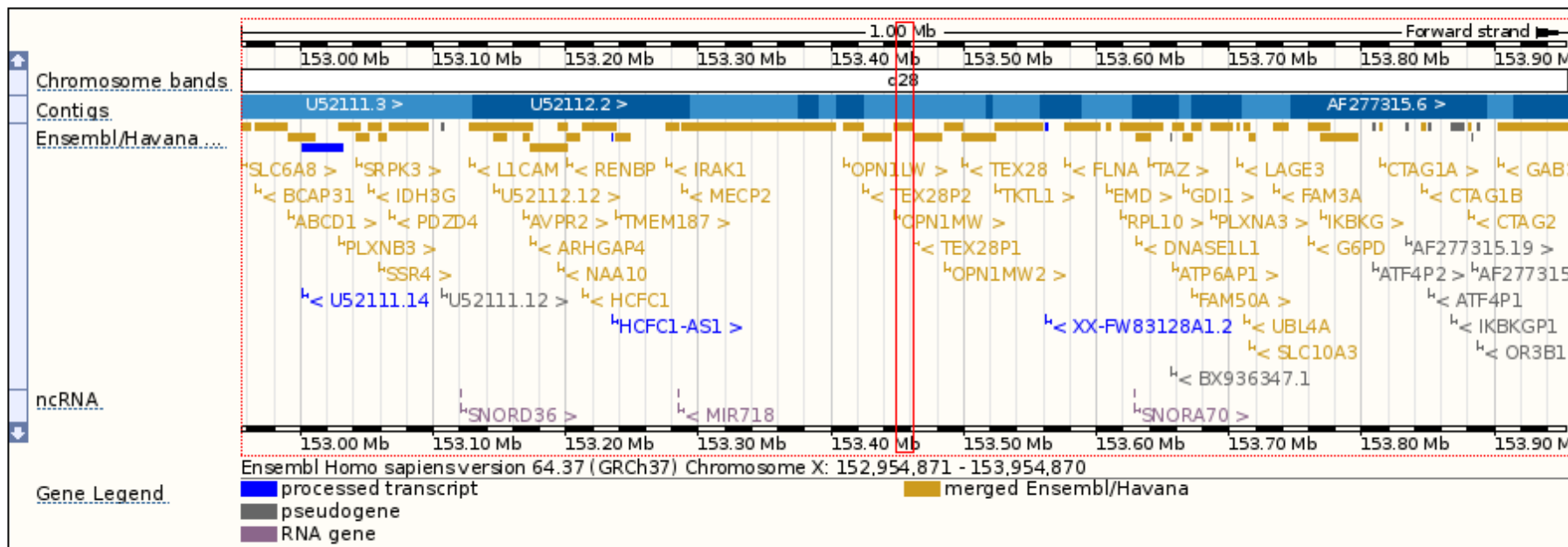
http://uswest.ensembl.org/Homo_sapiens/Location/View?h=Havana%20gene;r=X

Chromosome X: 153,448,109-153,461,632



[Export Image](#)

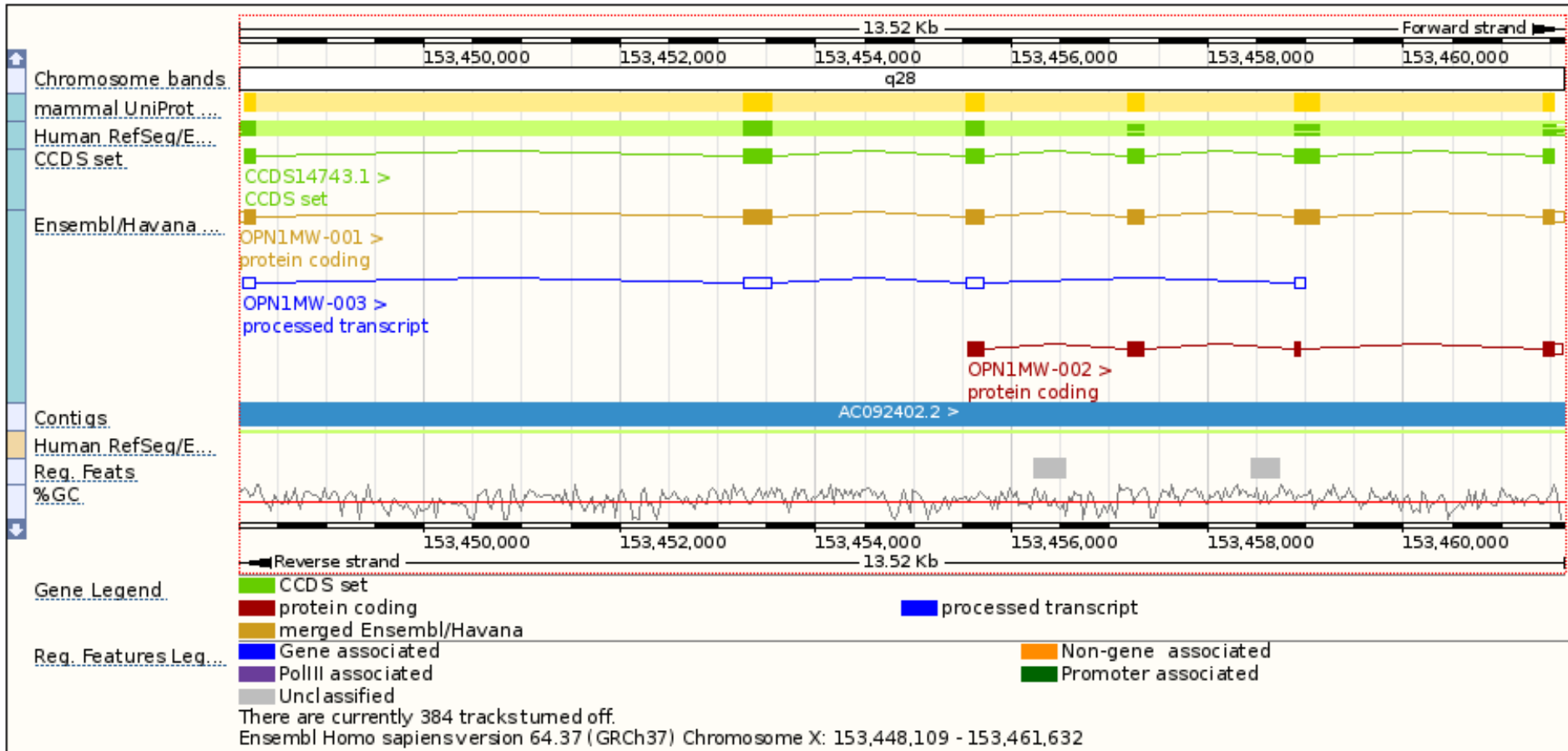
Region in detail [help](#)



[Export Image](#)

Ensembl OPN1MW Transcripts

http://uswest.ensembl.org/Homo_sapiens/Location/View?h=Havana%20gene;r=X



[Export Image](#)



Ensembl Tutorials

<http://uswest.ensembl.org/info/website/tutorials/index.html>



Ensembl Tutorials and Worked Examples

Online tutorials

- [Introduction to Browsing Chordate Genomes](#) - a 3-hour comprehensive interactive tutorial aimed at beginners

Video tutorials

The tutorials listed below are Flash animations of some of our training presentations. We are gradually adding to the list, so please check back regularly.



Note that we are now hosting all our tutorials on [YouTube](#) (and [优酷网](#) for users in China) for ease of maintenance. A selection of tutorials is also available on the [EBI E-Video website](#).

Title	Running time (minutes)
The Ensembl Genome Browser	10:00
EnsemblGenomes - Extending Ensembl across the taxonomic space	4:54
Comparative Genomics	9:58
SNPs and other Variations - 1 of 2	7:06
SNPs and other Variations - 2 of 2	6:30
Introduction to BioMart	4.27
BioMart: Variation IDs to HGNC Symbols	2:58
Clip: Exons and Introns	1:12
Clip: Export Sequence	1:08
Clip: Genome Variation	0:37
Clip: Transcriptomics (ArrayExpress)	1:12
Clip: View Conserved Sequence	1:26
Clip: View External Data (DAS)	1:53

[Genome Browser](#)[ENCODE](#)[Neandertal](#)[Blat](#)[Table Browser](#)[Gene Sorter](#)[In Silico PCR](#)[Genome Graphs](#)[Galaxy](#)[VisiGene](#)[Proteome Browser](#)[Utilities](#)[Downloads](#)[Release Log](#)[Custom Tracks](#)

About the UCSC Genome Bioinformatics Site

Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for a large collection of genomes. It also provides portals to the [ENCODE](#) and [Neandertal](#) projects.

We encourage you to explore these sequences with our tools. The [Genome Browser](#) zooms and scrolls over chromosomes, showing the work of annotators worldwide. The [Gene Sorter](#) shows expression, homology and other information on groups of genes that can be related in many ways. [Blat](#) quickly maps your sequence to the genome. The [Table Browser](#) provides convenient access to the underlying database. [VisiGene](#) lets you browse through a large collection of *in situ* mouse and frog images to examine expression patterns. [Genome Graphs](#) allows you to upload and display genome-wide data sets.

The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the Center for Biomolecular Science and Engineering ([CBSE](#)) at the University of California Santa Cruz ([UCSC](#)). If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our [public mailing list](#).

News

[News Archives](#) ►

To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, subscribe to the [genome-announce](#) mailing list.

10 October 2011 - Updated Yeast Browser Released

We are happy to announce an updated Yeast Genome Browser for the initial release of *Saccharomyces cerevisiae*, sacCer3. The April 2011 *Saccharomyces cerevisiae* genome assembly (Saccharomyces cerevisiae S288c assembly from Saccharomyces Genome Database (GCA_000146055.2)) was produced by the [Saccharomyces Genome Database \(SGD™\)](#) project.

Chromosomes available in this assembly: chrI, chrII, chrIII, chrIV ... etc ... chrXVI, and chrM. See also: [SGD™ genome snapshot/overview](#).

Downloads of the yeast data and annotations may be obtained from the UCSC Genome Browser [FTP server](#) or [Downloads](#) page. The *S. cerevisiae* annotation tracks were generated by UCSC and collaborators worldwide.

We'd like to thank the Saccharomyces Genome Database (SGD™). The *S. cerevisiae* Genome Browser and annotation tracks were produced by Hiram Clawson, Greg Roe, and Steve Heitner. See the [Credits](#) page for a detailed list of the organizations and individuals who contributed to this release.

Human (*Homo sapiens*) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

clade	genome	assembly	position or search term	gene
Mammal ▾	Human ▾	Feb. 2009 (GRCh37/hg19) ▾	opsin	<input type="text"/>
				submit
Click here to reset the browser user interface settings to their defaults.				
track search		add custom tracks		track hubs
configure tracks and display			clear position	

About the Human Feb. 2009 (GRCh37/hg19) assembly ([sequences](#))

The February 2009 human reference sequence (GRCh37) was produced by the [Genome Reference Consortium](#).

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the [User's Guide](#) for more information.

Request:**Genome Browser Response:**

chr7	Displays all of chromosome 7
chrUn_gl000212	Displays all of the unplaced contig gl000212
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm telomere
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting with position 1000000
RH18061;RH80175	Displays region between genome landmarks, such as the STS markers RH18061 and RH80175. This syntax may also be used for other range queries, such as between uniquely determined ESTs, mRNAs, refSeqs, etc.



Homo sapiens
(Graphic courtesy of [CBSE](#))

[OPN4 \(uc010gmk.1\) at chr10:88414314-88426605](#) - opsin 4 isoform 2
[OPN4 \(uc009xsx.1\) at chr10:88420233-88423870](#) - opsin 4 isoform 1
[OPN1MW \(uc004fkd.2\) at chrX:153485203-153499469](#) - opsin 1 (cone pigments), medium-wave-sensitive
[OPN1MW \(uc004fkb.2\) at chrX:153448085-153462351](#) - opsin 1 (cone pigments), medium-wave-sensitive
[OPN1LW \(uc004fjz.3\) at chrX:153409725-153424505](#) - opsin 1 (cone pigments), long-wave-sensitive
[OPN1SW \(uc003vnt.3\) at chr7:128412545-128415844](#) - opsin 1 (cone pigments), short-wave-sensitive
[OPN5 \(uc003ozd.2\) at chr6:47754862-47794114](#) - opsin 5 isoform 2
[OPN5 \(uc003ozc.2\) at chr6:47749798-47794114](#) - opsin 5 isoform 1
[OPN4 \(uc001kdg.2\) at chr10:88414314-88426214](#) - opsin 4 isoform 1
[OPN4 \(uc001kdp.2\) at chr10:88414314-88426214](#) - opsin 4 isoform 2
[OPN3 \(uc001hza.2\) at chr1:241756453-241803701](#) - opsin 3
[RGR \(uc001kde.1\) at chr10:86004809-86018944](#) - retinal G-protein coupled receptor isoform 3
[RGR \(uc001kdd.1\) at chr10:86004809-86018944](#) - retinal G-protein coupled receptor isoform 1
[RGR \(uc001kdc.1\) at chr10:86004809-86018944](#) - retinal G-protein coupled receptor isoform 2
[VSX1 \(uc010gde.1\) at chr20:25052132-25062767](#) - visual system homeobox 1 isoform b
[VSX1 \(uc002wue.2\) at chr20:25052132-25062767](#) - visual system homeobox 1 isoform b
[VSX1 \(uc002wug.1\) at chr20:25059195-25062767](#) - visual system homeobox 1 isoform b
[VSX1 \(uc010gdf.1\) at chr20:25052132-25062767](#) - visual system homeobox 1 isoform b
[VSX1 \(uc010gdd.1\) at chr20:25052132-25062767](#) - visual system homeobox 1 isoform b
[VSX1 \(uc002wuf.2\) at chr20:25056100-25062767](#) - visual system homeobox 1 isoform a
[CRX \(uc002phg.3\) at chr19:48325099-48346584](#) - cone-rod homeobox protein
[RHO \(uc003emt.2\) at chr3:129247482-129254186](#) - rhodopsin
[RRH \(uc003hzv.2\) at chr4:110749150-110765859](#) - peropsin
[ARR3 \(uc004dyb.2\) at chrX:69488185-69501690](#) - arrestin 3, retinal (X-arrestin)
[GRK7 \(uc011bnd.1\) at chr3:141497043-141535890](#) - G-protein-coupled receptor kinase 7 precursor

RefSeq Genes

[OPN1MW at chrX:153485203-153499470](#) - (NM_000513) medium-wave-sensitive opsin 1
[OPN1MW at chrX:153448085-153462352](#) - (NM_000513) medium-wave-sensitive opsin 1
[OPN5 at chr6:47749775-47794116](#) - (NM_181744) opsin-5
[OPN1SW at chr7:128412543-128415844](#) - (NM_001708) short-wave-sensitive opsin 1
[OPN1LW at chrX:153409725-153424507](#) - (NM_020061) long-wave-sensitive opsin 1
[OPN1MW2 at chrX:153485203-153498755](#) - (NM_001048181) medium-wave-sensitive opsin 1
[OPN1MW2 at chrX:153448085-153461637](#) - (NM_001048181) medium-wave-sensitive opsin 1
[OPN3 at chr1:241756452-241803701](#) - (NM_014322) opsin-3

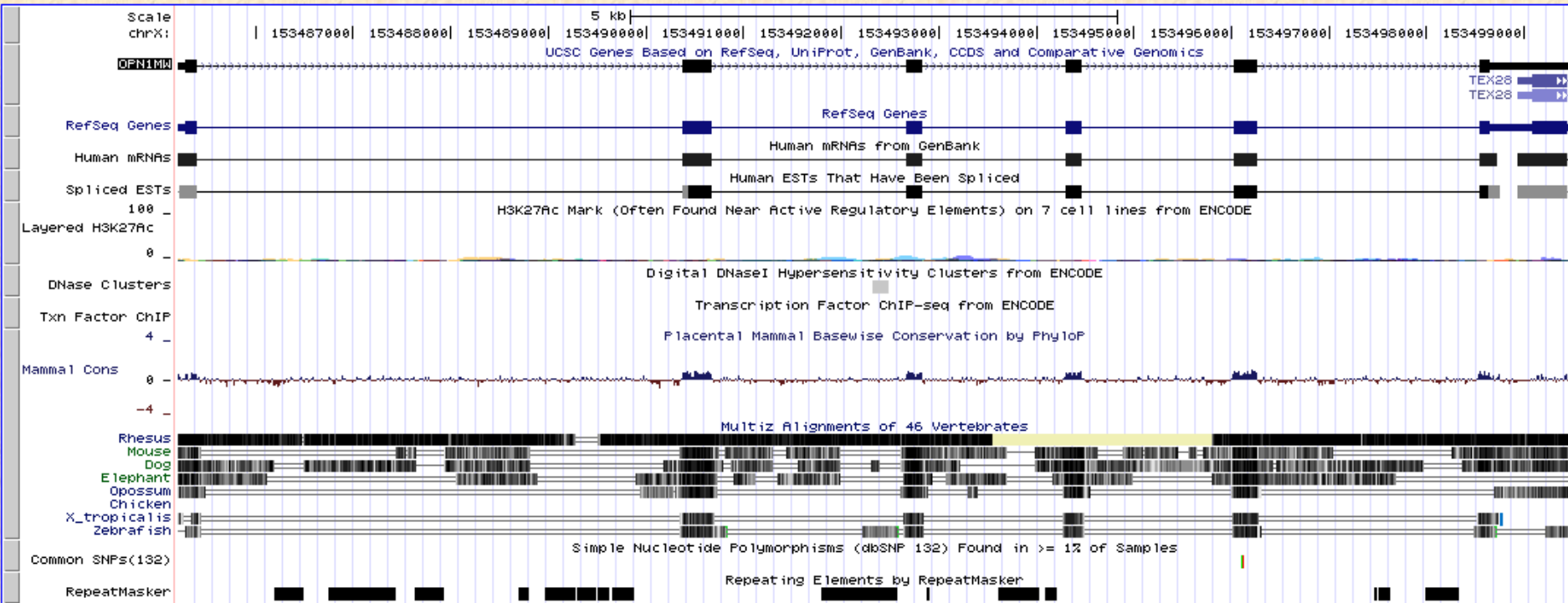
Non-Human RefSeq Genes

[OPN5 at chr6:47749803-47779473](#) - (NM_001206080) opsin-5
[OPN1SW at chr7:128413703-128415835](#) - (NM_001076704) blue-sensitive opsin
[LOC751972 at chr3:129247577-129252528](#) - (NM_001076696) RH2 opsin

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chrX:153,485,203-153,499,469: jump clear size 14,267 bp. configure



move start < 2.0 >

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

move end < 2.0 >



http

collapse all expand all
 Use drop-down controls below and press refresh to alter tracks displayed.
 Tracks with lots of items will automatically be displayed in more compact modes.

Mapping and Sequencing Tracks refresh

Base Position dense ▾	Chromosome Band hide ▾	STS Markers hide ▾	18 FISH Clones hide ▾	Recomb Rate hide ▾	Map Contigs hide ▾
Assembly hide ▾	GRC Map Contigs hide ▾	Gap hide ▾	BAC End Pairs hide ▾	18 Fosmid End Pairs hide ▾	GC Percent hide ▾
GRC Patch Release [No data-chrX]	Hg18 Diff hide ▾	NCBI Incident hide ▾	Short Match hide ▾	Restr Enzymes hide ▾	Wiki Track hide ▾
<input checked="" type="checkbox"/> BU ORChID hide ▾	<input checked="" type="checkbox"/> Mapability hide ▾				

983641&known

Phenotype and Disease Associations refresh

18 GAD View hide ▾	DECIPHER hide ▾	OMIM AV SNPs hide ▾	OMIM Genes hide ▾	OMIM Pheno Loci hide ▾	GWAS Catalog hide ▾
18 RGD Human QTL hide ▾	18 RGD Rat QTL hide ▾	18 MGI Mouse QTL hide ▾			

Genes and Gene Prediction Tracks refresh

UCSC Genes pack ▾	Alt Events hide ▾	<input checked="" type="checkbox"/> GENCODE Genes V4 hide ▾	<input checked="" type="checkbox"/> GENCODE Genes V7 hide ▾	CCDS hide ▾	RefSeq Genes dense ▾
Other RefSeq hide ▾	MGC Genes hide ▾	ORFeome Clones hide ▾	TransMap... hide ▾	Vega Genes hide ▾	Ensembl Genes hide ▾
AceView Genes hide ▾	N-SCAN hide ▾	SGP Genes hide ▾	Geneid Genes hide ▾	Genscan Genes hide ▾	Exoniphy hide ▾
tRNA Genes hide ▾	H-Inv 7.0 hide ▾	17 EvoFold hide ▾	sno/miRNA hide ▾	IKMC Genes Mapped hide ▾	

mRNA and EST Tracks refresh

Human mRNAs dense ▾	Spliced ESTs dense ▾	Human ESTs hide ▾	Other mRNAs hide ▾	Other ESTs hide ▾	18 H-Inv hide ▾
Gene Bounds hide ▾	18 SIB Alt-Splicing hide ▾	18 Poly(A) hide ▾	18 CGAP SAGE hide ▾	Human RNA Editing hide ▾	

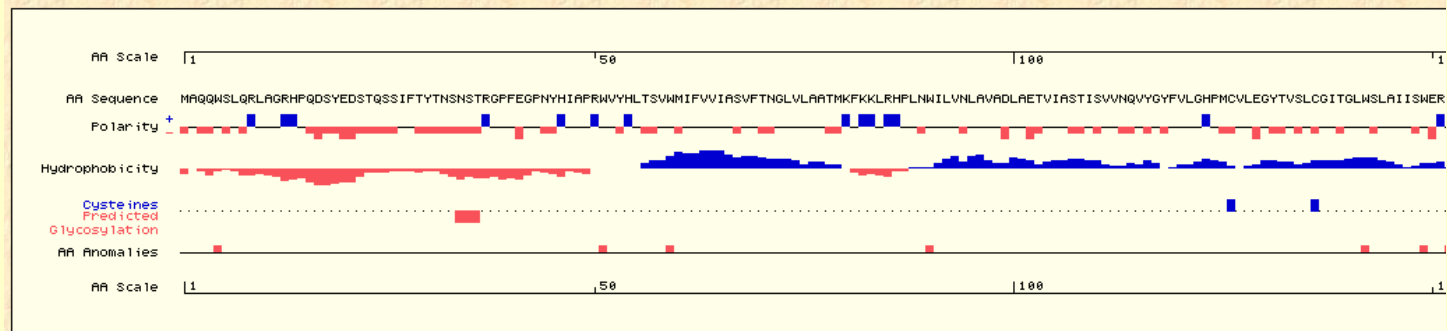
Expression refresh

Affy Exon Array hide ▾	Affy GNF1H hide ▾	<input checked="" type="checkbox"/> Affy RNA Loc hide ▾	Affy U133 hide ▾	Affy U133Plus2 hide ▾	Affy U95 hide ▾
Allen Brain hide ▾	Burge RNA-seq hide ▾	<input checked="" type="checkbox"/> ENC Exon Array... hide ▾	<input checked="" type="checkbox"/> ENC ProtGeno... hide ▾	<input checked="" type="checkbox"/> ENC RNA-seq... hide ▾	GNF Atlas 2 hide ▾
18 Illumina WG-6	18 Sestan Brain				

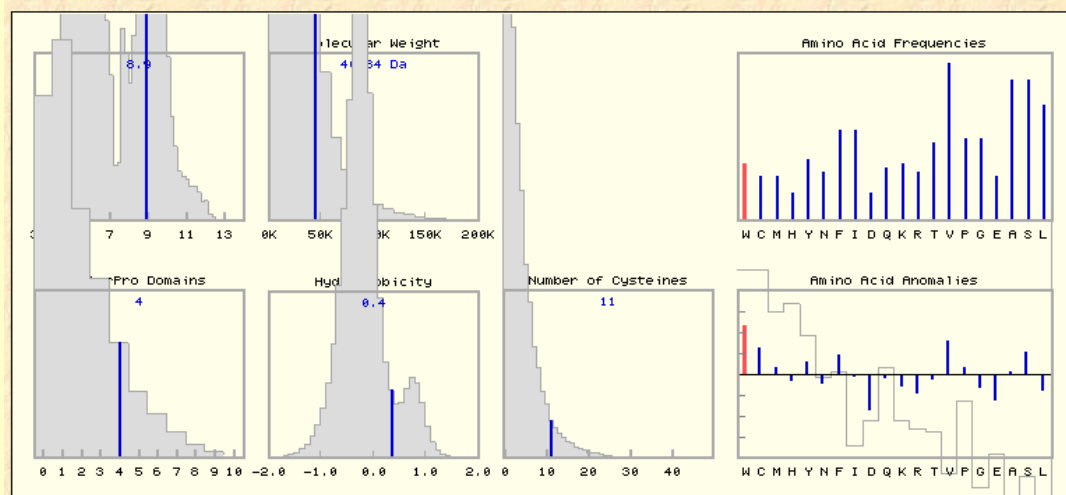
Protein **P04001** (aka **OPSG_HUMAN**) RecName: Full=Green-sensitive opsin; AltName: Full=Green cone photoreceptor pigment;

Organism: Homo sapiens (human)

Move [<<<](#) [<<](#) [<](#) [>](#) [>>](#) [>>>](#) Current scale: FULL Rescale to [1/6](#) [1/2](#) [FULL](#)



Explanation of Protein Tracks



Explanation of Protein Property Histograms

UCSC Links:

- [BLAT - Homo sapiens \(human\)](#)

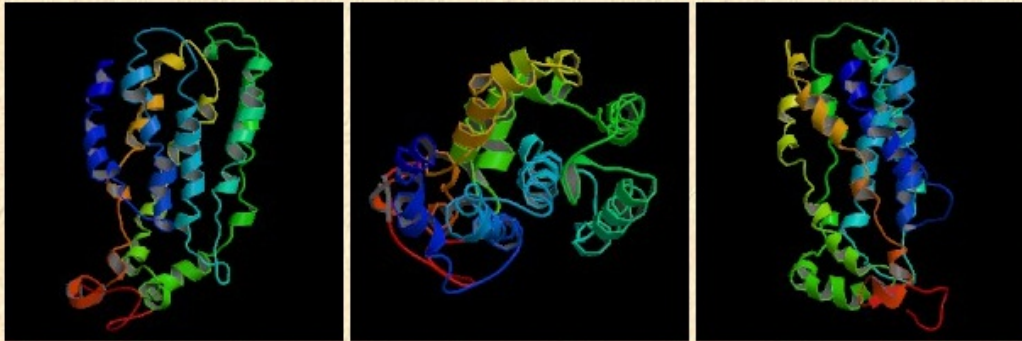
InterPro Domains: [Graphical view of domain structure](#)

- [IPR000276](#) - 7TM_GPCR_Rhodpsn
- [IPR017452](#) - GPCR_Rhodpsn_supfam
- [IPR001760](#) - Opsin
- [IPR000378](#) - Opsin_red/gm

Pfam Domains:

- [PF00001](#) - 7 transmembrane receptor (rhodopsin family)

Protein Data Bank (PDB) 3-D Structure

[1KPW](#) - ModelModBase Predicted Comparative 3D Structure on [P04001](#)

Front

Top

Side

The pictures above may be empty if there is no ModBase structure for the protein. The ModBase structure frequently covers just a fragment of the protein. You may be asked to log onto ModBase the first time you click on the pictures. It is simplest after logging in to just click on the picture again to get to the specific info on that model.

Total amino acids: 364

FASTA record:

```
>P04001|OPSG_HUMAN|RecName: Full=Green-sensitive opsin; AltName: Full=Green cone photoreceptor pigment;
MAQQWSLQRLAGRHPQDSYEDSTQSSIFTYTNSNSTRGPFEGPNYHIAPR
WVYHLTSVWMIFFVVIASVFTNGLVLAATMKFKKLRHPLNWILVNLAVADL
AETVIASTISVVNQVYGFVLGHPMCVLEGYTVSLCGITGLWSLAIISWE
RWMVVCCKPFGNVRFDKLAIVGIAFSWIWAAVWTAPPIFGWSRYWPHGLK
TSCGPDVFSGSSYPGVQSYMIVLMVTCITPLSIIIVLCYLQVWLAIRAVA
KQKBESESTQKAEKEVTRMVVVMVLAFCFCWGPYAFFACFAAANPGYPFH
PLMAALPAFFAKSATIYNPVIYVFMNRQFRNCILQLFGKKVDDGSELSSA
SKTEVSSVSSVSPA
```

UCSC Genome Bioinformatics

[Home](#) - [Genomes](#) - [Blat](#) - [Tables](#) - [Gene Sorter](#) - [PCR](#) - [VisiGene](#) - [Proteome](#) -
[Session](#) - [FAQ](#)

Genome Browser User Guide

Table of Contents:

- [What does the Genome Browser do?](#)
- [Getting started: Genome Browser gateways](#)
- [Fine-tuning the Genome Browser display](#)
- [Annotation track descriptions](#)
- [Using BLAT alignments](#)
- [Getting started on the Table Browser](#)
- [Getting started using Sessions](#)
- [Getting started on Genome Graphs](#)
- [Using the VisiGene Image Browser](#)
- [DNA text formatting](#)
- [Converting data between assemblies](#)
- [Downloading genome data](#)
- [Creating and Managing custom annotation tracks](#)

Search the Genome Browser help pages:

Search the Genome mailing list archives:

See also the [Open Helix tutorial and training materials](#).

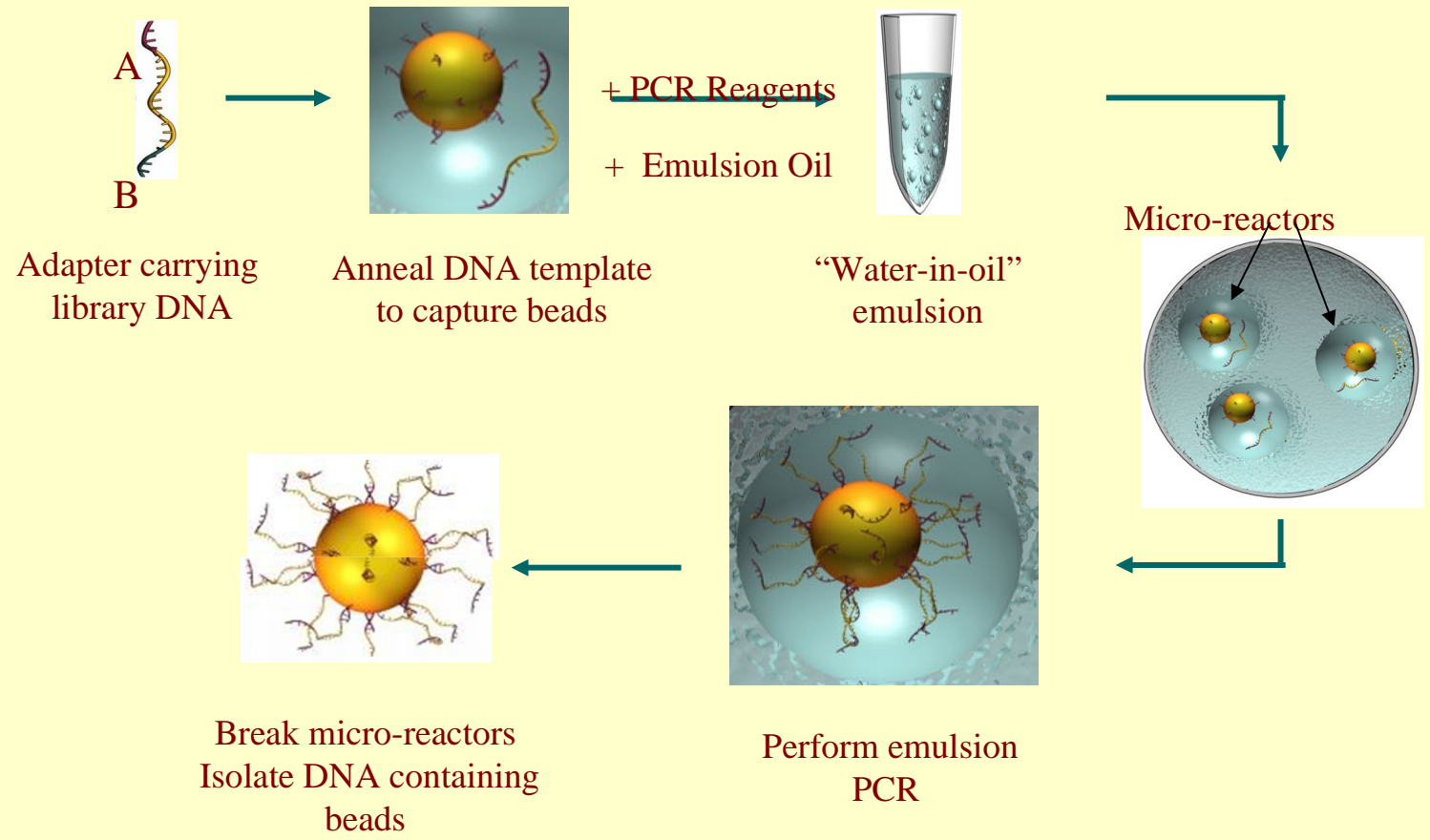
[Questions and feedback are welcome.](#)

What does the Genome Browser do?



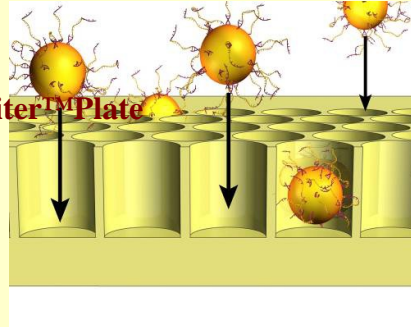
Emulsion Based Clonal Amplification

Single test tube generation of millions of clonally amplified sequencing templates
No cloning and colony picking

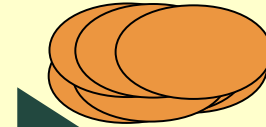


Depositing DNA Beads into the PicoTiter™ Plate

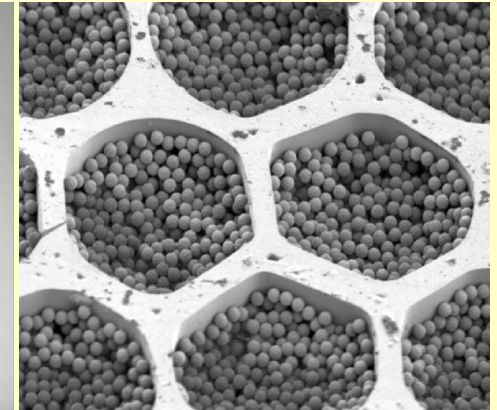
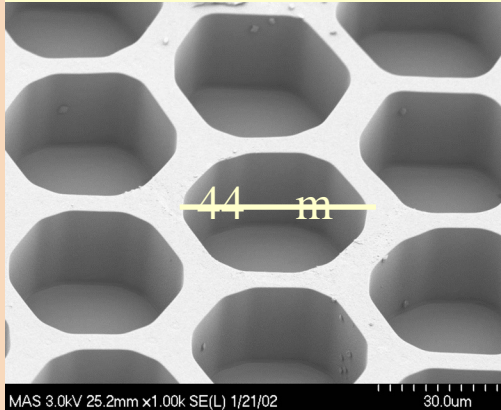
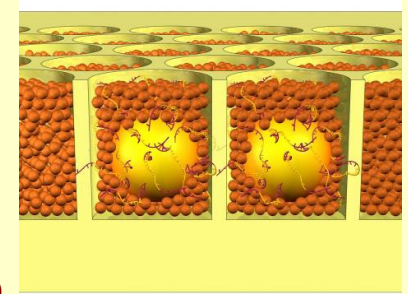
Load beads into PicoTiter™ Plate



Load Enzyme Beads



Centrifuge Step



- 70x75mm array of fused optical fibers with etched wells
- 1.6 million wells monitored optically through fiber

Sequencing-By-Synthesis

- Simultaneous sequencing of the entire genome in hundreds of thousands of picoliter-size wells
- Pyrophosphate signal generation

