

Genetic Variation

Zygosity in diploid organisms

- Homozygous: two identical copies of an allele
- Heterozygous: two non-identical copies of an allele
- Hemizygous: only one copy of allele is present

Genetic Variation

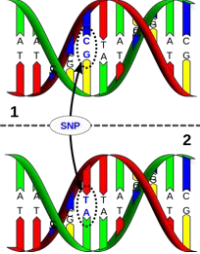
Allele: one of multiple possible forms of a genetic locus.

- One or more alleles affect a particular trait.
- Different alleles of a gene often result in the same trait.
- Humans and most other animals have two identical or different alleles of each gene (one per sister chromatid).
- Dominant allele: acts dominantly over a recessive allele.
- Recessive allele: acts recessively to a dominant allele.
- Can also have codominance and incomplete dominance.

Genetic Variation

Indels (insertions and deletions), single nucleotide polymorphisms (SNPs), and Copy Number variants (CNVs)

- >10 million known SNPs in humans.
- Genetic variation important in disease and research.




~99% DNA similarity on average amongst individuals

Image: MIT, Christine Daniloff

Genetic Variation

	Point mutations				
	No mutation	Silent	Nonsense	Missense	
				conservative	non-conservative
DNA level	TTC	TTT	ATC	TCC	TGC
mRNA level	AAG	AAA	UAG	AGG	ACG
protein level	Lys	Lys	STOP	Arg	Thr



- Silent mutations:** don't change the amino acid sequence.
- Nonsense mutations:** encode stop codons.
- Missense mutations:** cause a change in the amino acid sequence.

Assessing Genetic Variation

- Genome sequencing, targeted or total
- PCR – example from class

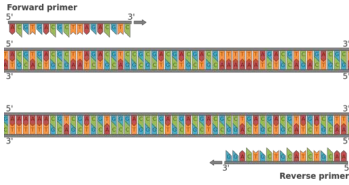
```

CTAAAGATGATCTTAGTCCCGGTTCGAA
TCTTTAGTCCCAGTTGATACACCAAGC
GTAATACCAACCCGGACTAAGATCCCG
GGACTAAAGTCCCAGCCCTATATATG
TTCAAAATTTCTTCAAAAAGAGGGGAG
GTGATTACATACAATCGGAGGTGCCTA
TTTGTATAC TACAATTGCACCTATGTTT
GTAAGTTGATGAGAGAGAAAATGTGTGT
TTTGCTAAACAAGGTTTTATAAATAGTTC
AAATAATAGAAAACAACATAAATGAAAT
TATTACTTAACAATATGTTTTAAGAATAT
AAATAAGATATCTATAATATGTATGACT
ACGGTTTTTTTGACTCATGTAGATGGATC
    
```

Primers and Oligonucleotides

Oligonucleotides: short DNA or RNA molecules.

Primers: DNA oligonucleotides used in PCR. Specify target DNA sequence and facilitate amplification.



Primers and Oligonucleotides

Primer optimization: GC content, T_m, length, specificity.

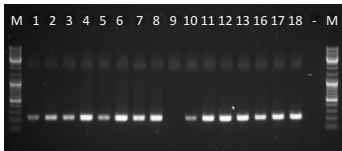
Go to:

<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>

or google:

primer3 plus cgi

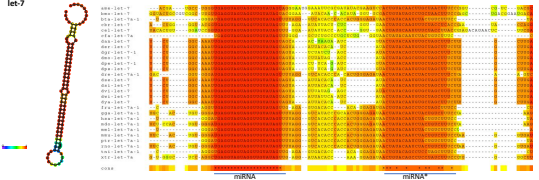
PCR Products Sent for Sequencing



Sequence Alignments

Pairwise alignment: alignment of two sequences to identify the best match.
Multiple sequence alignment: alignment of three or more sequences.


let-7





The figure shows a secondary structure diagram of the let-7 miRNA on the left, represented as a stem-loop structure with colored nucleotides. To its right is a heatmap representing a multiple sequence alignment of let-7 miRNAs from various species. The heatmap uses a color scale from red (high conservation) to yellow (low conservation) to show conserved regions across the sequences. The y-axis lists various let-7 family members like let-7a-1, let-7a-2, etc., and the x-axis represents the sequence positions.

Sequence Alignments

Some common alignment software:

TCoffee: software for progressive alignment of two or more DNA, RNA, or protein sequences. 

Clustal Omega: software for global alignment of multiple DNA, RNA, or protein sequences (up to several hundred thousand). 

Basic Local Alignment Search Tool (BLAST): a popular search tool for comparing DNA or amino acid sequences. 

BLAST

Query sequence is broken down into short fragments ("words") of 3 aa or 28 nt (by default)

↓

Words are matched to a database of words

↓

Two non-overlapping word matches within a short distance of each other are identified

↓

Maximal alignment identified by extending from better matching word in both directions

BLAST

Program	Query	Subject
blastn	DNA	DNA
blastp	Protein	Protein
blastx	DNA translated to protein	Protein
tblastn	Protein	DNA translated to protein
tblastx	DNA translated to protein	DNA translated to protein
blast2	DNA or protein	DNA or protein

Several databases to blast against:

- nr – all non-redundant sequences in GenBank
- Species-specific sequences (e.g. human, mouse, etc.)
- Custom databases
- Use on web, server, or personal computer

Exercise 2:

- Examine the sequence trace from your sequencing run.
- Blast your sequence and determine what gene it is.
- Determine the function of the gene.
- Align your sequence to the wild type and mutant reference sequence.
- Answer the following questions and submit your answers on Canvas:
 1. What is the identity of the gene?
 2. What trait is the gene linked to?
 3. Do you contain any of the common variants that are associated with the trait?
