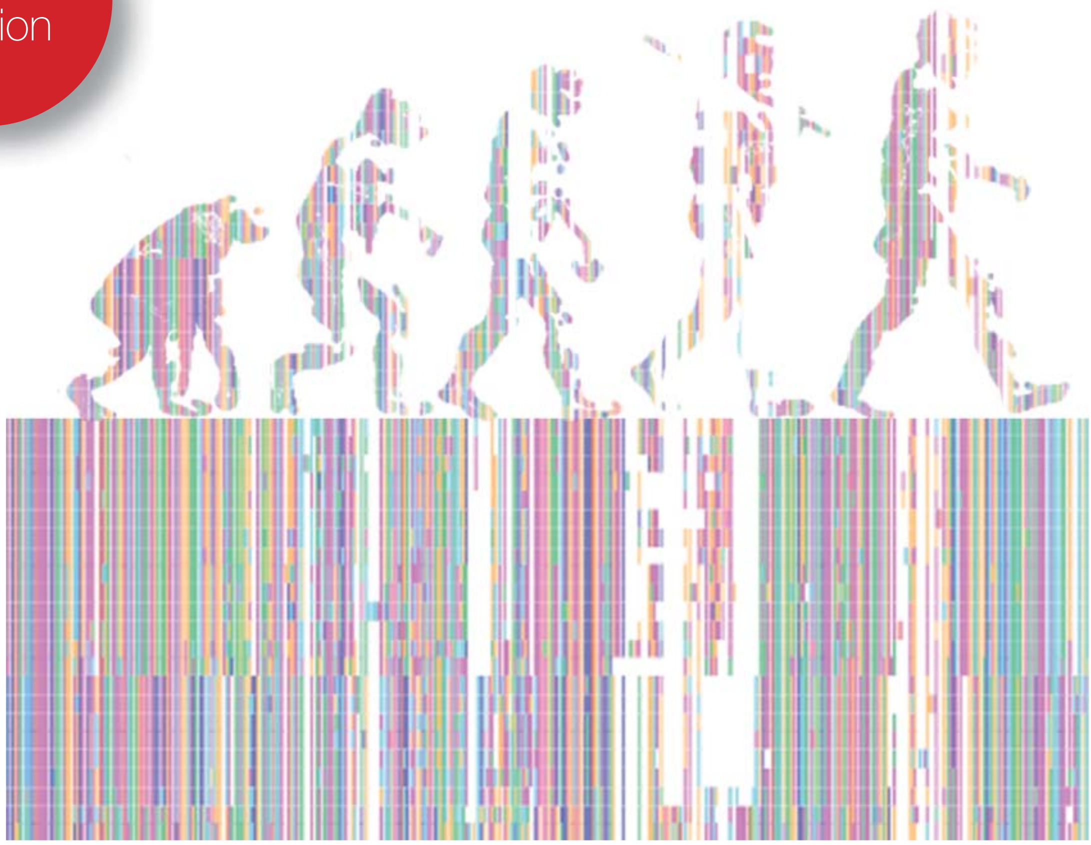
A First Look at the Code of Life

Cleopatra Kozlowski

An introductory course to concepts in Bioinformatics

English version



ELLS – European Learning Laboratory for the Life Sciences

EMBL

Cleopatra Kozlowski

A First Look at the Code of Life

Student Handout

Table of contents

1 Gene Finding

1.1 The 20 Amino Acids and their Symbols

Name	Abbreviations	
Alanine	ala	a
Arginine	arg	r
Asparagine	asn	n
Aspartic acid	asp	d
Cysteine	cys	c
Glutamine	gln	q
Glutamic acid	glu	e
Glycine	gly	g
Histidine	his	h
Isoleucine	ile	i
Leucine	leu	l
Lysine	lys	k
Methionine	met	m
Phenylalanine	phe	f
Proline	pro	p
Serine	ser	s
Threonine	thr	t
Tryptophan	trp	w
Tyrosine	tyr	y
Valine	val	v

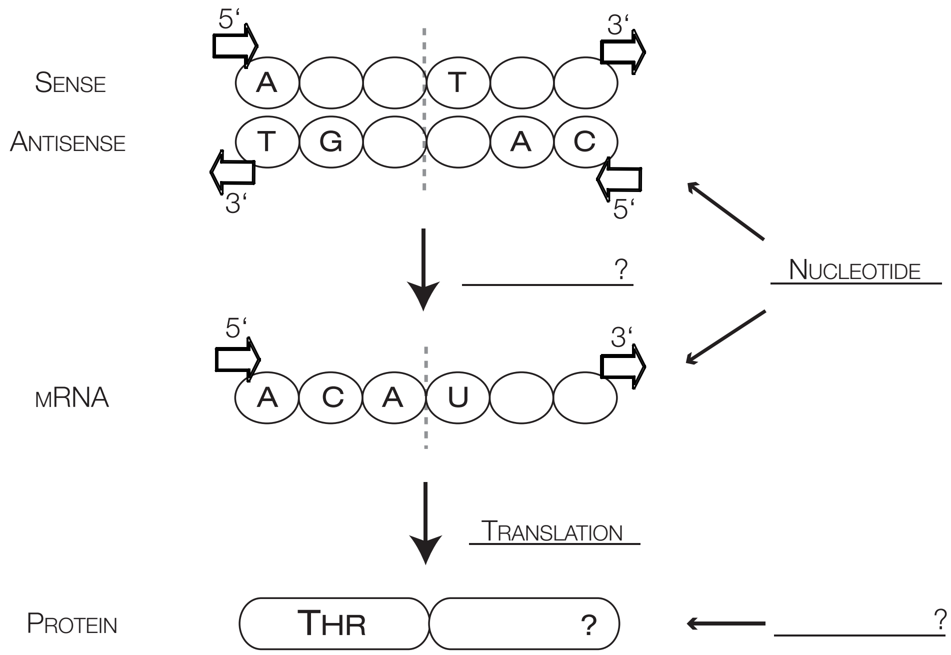
1.2 Table of the Genetic Code

In the presence of zinc ions, insulin forms hexamers (groups of 6 molecules, Fig. 1.9), resulting in a torus-like (or "doughnut") shape. Insulin is stored in β-cells and secreted in the bloodstream as a hexamer. However, the active form is a monomer.

1st Base	2nd Base				3rd Base
	U	C	A	G	
U	UUUP Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine	U
	UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C
	UUA Leucine	UCA Serine	UAA Stop	UGA Stop	A
	UUG Leucine	UCG Serine	UAG Stop	UGG Tryptophan	G
C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U
	CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine	C
	CUA Leucine	CCA Proline	CAA Glutamine	CGA Arginine	A
	CUG Leucine	CCG Proline	CAG Glutamine	CGG Arginine	G
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Arginine	C
	AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine	A
	AUG Methionine	ACG Threonine	AAG Lysine	AGG Arginine	G
G	GUU Valine	GCU Alanine	GAU Aspartic Acid	GGU Glycine	U
	GUC Valine	GCC Alanine	GAC Aspartic Acid	GGC Glycine	C
	GUA Valine	GCA Alanine	GAA Glutamic Acid	GGA Glycine	A
	GUG Valine	GCG Alanine	GAG Glutamic Acid	GGG Glycine	G

1.3 Exercise

Fill in the missing parts of the diagram.



1.4 Gene Finding Game

Find the following short peptides in the DNA sequence given below.

- Met Ser Ile Leu Leu Tyr Stop
- Met Ser Ile Leu Val Glu Arg Stop
- Met Cys Arg Thr Stop

1-50	TGGTCCTGCA GTCCTCTCCT GGCGCCCCGG GGGCGAGCGG ATGTCGATTC
51-100	TCGTGGAAAG ATAGTCCCGC TGCCTGCGGG CGGAGGGACC GTGCTGACCA
101-150	AGATGTACCC GCGCGGCAAC CACTGGGCGG TGGGGCACTT AATGGGGAAA
151-200	AAGAGCACAG GGGAGTCTTC TTCTGTATGT TCTGAGAGAG GGAGCCTGAA
201-250	GCAGCAGCTG AGAGATGTGT CGAACGTGAA GTACATCAGG TGGGAAGAAG
251-300	CTGCAAGGAA TTTGCTGGGT CTCATAGAAG CTAAAGGAGA ACAGAAACCA
301-350	CCAGCCACCT CAACCCAAGA TGTCGATTCT ACTTTATTAA GCCCTGGGCA
351-400	ATCAGCAGCC TTCGTGGGAT ATGTCAGAGG ATAGCAGCCA ACTGAATAGC

1.5 Questions

1. Which amino acid does AGG code for?

2. Which codons code for Serine?

3. How many ways can ‘Cys Arg Thr’ be coded?

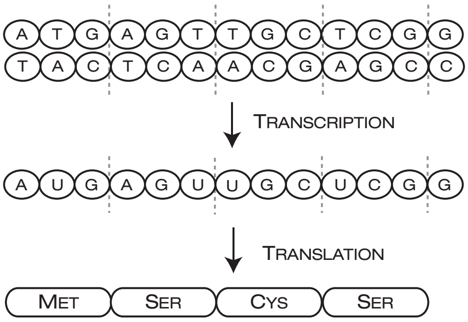
4. What is the ‘antisense’ strand of DNA?

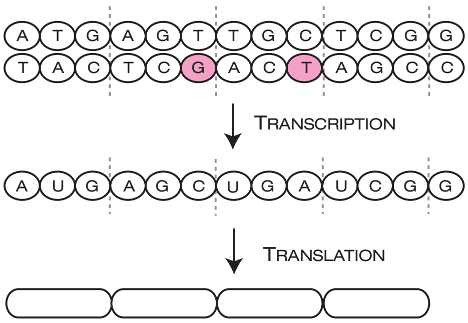
5. If you look at the table of codons, you see that the last nucleotide in the codon often doesn’t matter, for what amino acid is coded. What consequences may this have?

2 Mutations

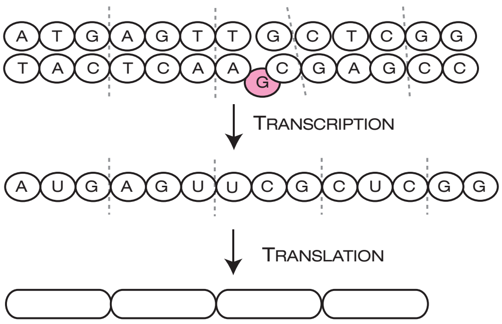
2.1 Exercise

Fill in the blanks in the diagram below.

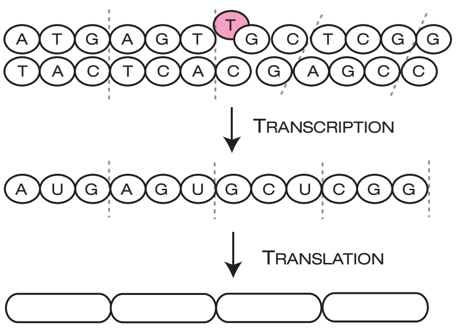




WHAT KIND OF MUTATION?
RESULT:



WHAT KIND OF MUTATION?
RESULT:



WHAT KIND OF MUTATION?
RESULT:

2.2 Questions

1. What kind of mutations will have the greatest effect on the organism?

2. How does the redundancy of the genetic code reduce the effects of mutations?

2.3 Table for Gene Finding Game

Use this table to fill in the scores for Lesson 2.

To play the game, use the cards in Appendix I and II.

NAME	TYPE OF MUTATION	SCORE

3 Phylogenetic Tree

3.1 Questions

1. What regions of DNA should you use to compare organisms that are closely related?

2. What kind of genes should you use to compare organisms which are evolutionarily distant from each other?

3. What should you do if you are comparing two sequences, but one of them has gaps, due to insertions/ deletions?

4. Can you think of reasons why this method of simply comparing the number of differences between the nucleotides cannot work, if you are comparing organisms that are very different? Think about the fact that we are assuming it takes 20 million years for every nucleotide in a sequence to mutate.

5. Can you think of other reasons why it may not be so good to use this method to calculate genetic distances? What simplifications have we made?

6. Can you think of reasons why if you are studying more distant organisms, it is better to compare amino acid sequences than DNA sequences?

4 Mobile DNA

4.1 Transposon Game

Delicious	and (1)	Fun	Christmas	Cake	Recipe
Preheat	the	oven	to	300	Degrees.
Cream	together	the	butter	and (2)	sugar.
Add	Eggs,	Flour	and (3)	spices	mixing
well.	Fill	buttered	pan	with	batter,
and (4)	bake	for	about	1	hour.

4.2 Questions

1. Have the 'transposons' destroyed the meaning of sentences, or added something new?

2. Do you think you could get a new kind of cake this way? What is the parallel to evolution?

3. What kind of problems could transposition cause when comparing sequences? Answer: If there is more than one copy of similar genes in the organisms you are comparing, you will not know which you should compare to construct the tree.

Appendix I: Amino Acid Cards

Cut along the black lines to use the amino acid cards for Lesson 2.

Ala Alanine GCU, GCC, GCA, GCG **+1**	**Arg** Arginine CGU, CGC, CGA, CGG, AGA, AGG **+1**
Cys Cysteine UGU, UGC **+3**	**Glu** Glutamic Acid GAA, GAG **+2**
Ile Isoleucine AUU, AUC, AUA **+1**	**Leu** Leucine UUG, UUA, CUU, CUC, CUA, CUG +1
Asn Asparagine AAU, AAC **+2**	**Ser** Serine UCU, UCC, UCA, UCG, AGU, AGC **+1**
Gly Glycine GGU, GGC, GGA, GGG **+1**	**Val** Valine GUU, GUC, GUA, GUG **+1**
Lys Lysine AAA, AAG **+2**	**Met** Methionine (start codon) AUG **+3**

Phe Phenylalanine UUU, UUC **+2**	**Pro** Proline CCU, CCC, CCA, CCG **+1**
Trp Tryptophan UGG **+3**	**Tyr** Tyrosine UAU, UAC **+2**
Asp Aspartic Acid GAU, GAC **+2**	**His** Histidine CAU, CAC **+2**
End Terminator UAA, UAG, UGA **+1**	**Thr** Threonine ACU, ACC, ACA, ACG **+1**
A U G **(Start)**	

Appendix II: Nucleic Acid Cards

Cut along the black lines to use the amino acid cards for Lesson 2.

A	G	U	C	A	G
U	C	A	G	U	C
A	G	U	C	A	G
U	C	A	G	U	C
A	G	U	C	A	G
U	C	A	G	U	C
A	G	U	C	A	G
U	C	A	G	U	C
A	G	U	C	A	G
U	C	A	G	U	C

Acknowledgements

For suggestions offered during the preparation of this activity, and for critical reading of the draft, I thank Alexandra Manaia and Julia Willingale-Theune.

The cover image by Petra Riedinger;

The Indo-European language tree and phylogenetic trees by Nicola Graf;

Layout design by Nicola Graf;

Edited by Corinne Kox.