



**Estimating ds and N**

Consider two aligned horizontal sequences

	<u>Site 1</u>	<u>Site 2</u>	<u>Site 3</u>	<u>Site 4</u>	<u>Site 5</u>	<u>Site 6</u>
<b>DNA</b>	ACA	ATA	AT <b>C</b>	TTT	AAT <b>T</b>	CAA
<b>AA</b>	T	I	<b>I</b>	F	N	Q
<b>DNA</b>	ACA	ATA	AC <b>C</b>	TTT	AAC <b>C</b>	CAA
<b>AA</b>	T	I	<b>T</b>	F	N	Q

Can one claim that  $dN/dS = 1$ , because there is one synonymous and one non-synonymous substitution?

17

# Estimating dS and dN

Consider two **aligned homologous** sequences

	<u>Site 1</u>	<u>Site 2</u>	<u>Site 3</u>	<u>Site 4</u>	<u>Site 5</u>	<u>Site 6</u>
<b>DNA</b>	ACA	ATA	ATC	TTT	AAT	CAA
<b>AA</b>	T	I	I	F	N	Q
<b>DNA</b>	ACA	ATA	ACC	TTT	AAC	CAA
<b>AA</b>	T	I	T	F	N	Q

Can one claim that  $dN/dS = 1$ , because there is one synonymous and one non-synonymous substitution?

# Universal genetic code

This genetic code has 61 sense (non-termination) codons

Substitution types

	Synonymous			Non-synonymous			To a stop codon
	Transitions	Transversions	Total	Transitions	Transversions	Total	Total
1st position:	8	0	8	140	26	166	9
2nd position:	0	0	0	148	28	176	7
3rd position:	58	68	126	2	48	50	7
-----							
Total	66	68	134	290	102	392	23

- Approximately 3:1 (392 N : 134 S) ratio when mutations are generated and **fixed** completely at random
- Non-random distribution over codon positions
  - **All** second position mutations are non-synonymous
  - **Most** (but not all) synonymous mutations are confined to the third position