

Coding sequence of a gene

Counts

... ATG|CGC|**A**TC|CGA|TGC|CGA|TGG|CCG|TAG
AGC|TTA|GCC|CGG|**T**TA|TAG|CCC|AAG|GCT
CG|CGA|TTA|GCG|CTA|TAC|GGG|AGG|TAA ...



Observed dN = **2**
Observed dS = **2**

*Take into account the 3-nucleotide context
of every observed mutation*

... ATG|CGC|**A**TC|CGA|TG**C**|CGA|TGG|**C**CG|TAG
AGC|**T**TA|**G**CC|CGG|**T**TA|TAG|**C**CC|AAG|GCT
CG|CGA|**T**TA|GCG|CTA|TAC|GGG|A**G**G|TAA ...



Permuted dN = **8**
Permuted dS = **3**

*Derive the expected number of mutations
with the same 3-nucleotide context*



Non-synonymous



Synonymous