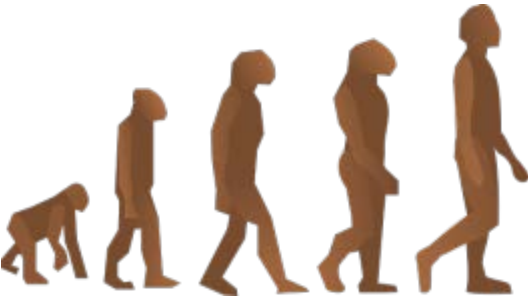# Searching for natural selection on individual genes: dN/dS
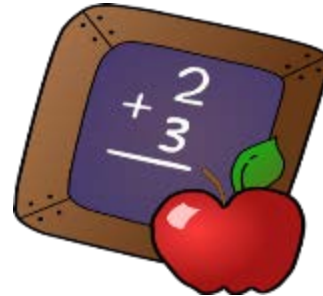
The fix: scale using number of synonymous changes!

- Synonymous differences accumulate neutrally
  - Can use them to scale for mutation rate differences

- RATIO of nonsynonymous to synonymous differences (**dN/dS**) estimates non-neutral changes relative to neutral changes

  - Let's see how this is done…

dN/dS Ratios

- dN = number of non-synonymous changes **per non-synonymous site**

- dS = number of synonymous changes **per synonymous site**

Calculating dN and dS

```
         ACT CCG GGG CCC            Thr Pro Gly Pro
          *       *   *
         ACG CCG GGC CTC            Thr Pro Gly Leu


site #   123 456 789 111
                     012

syn      001 001 001 001

non      110 110 110 110
```

Slide courtesy of M Rausher

Calculating dN and dS

```
ACT CCG GGG CCC          Thr Pro Gly Pro
 *          *   *
ACG CCG GGC CTC          Thr Pro Gly Leu
```

```
site # 123 456 789 111
                    012
```

Any change from A
gives different
amino acid

⋮ non-syn site

```
syn    001 001 001 001
```

```
non    110 110 110 110
```

Slide courtesy of M Rausher

Calculating dN and dS

```
        ACT CCG GGG CCC          Thr Pro Gly Pro
         *         *   *
        ACG CCG GGC CTC          Thr Pro Gly Leu

site #  123 456 789 111
                        012      ┌─────────────────────┐
                                 │   Any change in 2ⁿᵈ  │
syn     001 001 001 001          │   position gives     │
                                 │ different amino acid │
                                 │                      │
non     110 110 110 110          │  ⋮  non-syn site     │
                                 └─────────────────────┘
```

Any change in 2nd position gives different amino acid

non-syn site

Calculating dN and dS

```
ACT CCG GGG CCC
 *           *   *
ACG CCG GGC CTC
```

```
site # 123 456 789 111
                    012
```

```
syn    001 001 001 001
```

```
non    110 110 110 110
```

```
Thr Pro Gly Pro

Thr Pro Gly Leu
```

Any change DOES
NOT affect amino
acid

•
•• • syn site w/change

Slide courtesy of M Rausher

Calculating dN and dS

```
      ACT CCG GGG CCC              Thr Pro Gly Pro
       *         *   *
      ACG CCG GGC CTC              Thr Pro Gly Leu
```

```
site # 123 456 789 111            Total synonymous sites = 4
                   012            Total nonsynonymous sites = 8
```

Slide courtesy of M Rausher

```
syn     001 001 001 001

non     110 110 110 110
```

Calculating dN and dS

- What about bases where SOME changes affect amino acid and SOME changes do not???

  TTT = Phe, TTC = Phe, TTG = Leu, TTA = Leu

Calculating dN and dS

- What about bases where SOME changes affect amino acid and SOME changes do not???

  TTT = Phe, TTC = Phe, TTG = Leu, TTA = Leu

- If start with TTG, 3rd position has
  1/3 probability synonymous
  2/3 probability nonsynonymous
  … add to both totals in those amounts!

Calculating dN and dS

```
        ACT CCG GGG CCC          Thr Pro Gly Pro
         *       *   *
        ACG CCG GGC CTC          Thr Pro Gly Leu

site # 123 456 789 111
                    012

syn     001 001 001 001

non     110 110 110 110
```

Total synonymous sites = 4
Total nonsynonymous sites = 8

Total synonymous changes=2
Total nonsynonymous changes=1

Calculating dN and dS

```
ACT CCG GGG CCC          Thr Pro Gly Pro
 *       *   *
ACG CCG GGC CTC          Thr Pro Gly Leu
```

```
site # 123 456 789 111
                    012
```

Total synonymous sites = 4
Total nonsynonymous sites = 8

```
syn    001 001 001 001
```

Total synonymous changes=2
Total nonsynonymous changes=1

```
non    110 110 110 110
```

dN = 1/8 = 0.125   dS = 2/4 = 0.5

Slide courtesy of M Rausher

Calculating dN and dS

```
ACT CCG GGG CCC        Thr Pro Gly Pro
 *       *   *

ACG CCG GGC CTC        Thr Pro Gly Leu
```

```
site # 123 456 789 111
                   012
```

Total synonymous sites = 4
Total nonsynonymous sites = 8

```
syn    001 001 001 001
```

Total synonymous changes=2
Total nonsynonymous changes=1

```
non    110 110 110 110
```

dN = 1/8 = 0.125   dS = 2/4 = 0.5

dN/dS = 0.125/0.5 = 0.25

What does a dN/dS value *mean*?

- Estimates how much "non-neutral" (nonsynonymous) evolution has happened relative to "neutral" (synonymous) evolution
  - If a gene is evolving neutrally, has dN/dS = 1
    - No selection on nonsynonymous changes (null hyp)

What does a dN/dS value *mean*?

- Estimates how much "non-neutral" (nonsynonymous) evolution has happened relative to "neutral" (synonymous) evolution
  - If a gene is evolving neutrally, has dN/dS = 1
    - No selection on nonsynonymous changes (null hyp)
  - If a gene has dN/dS < 1, then "constrained"
    - Nonsynonymous changes selected against

What does a dN/dS value *mean*?

- Estimates how much "non-neutral" (nonsynonymous) evolution has happened relative to "neutral" (synonymous) evolution

  – If a gene is evolving neutrally, has dN/dS = 1
  - No selection on nonsynonymous changes (null hyp)

  – If a gene has dN/dS < 1, then "constrained"
  - Nonsynonymous changes selected against

  – If a gene has dN/dS > 1, then changing rapidly
  - Multiple nonsynonymous changes favored by natural selection
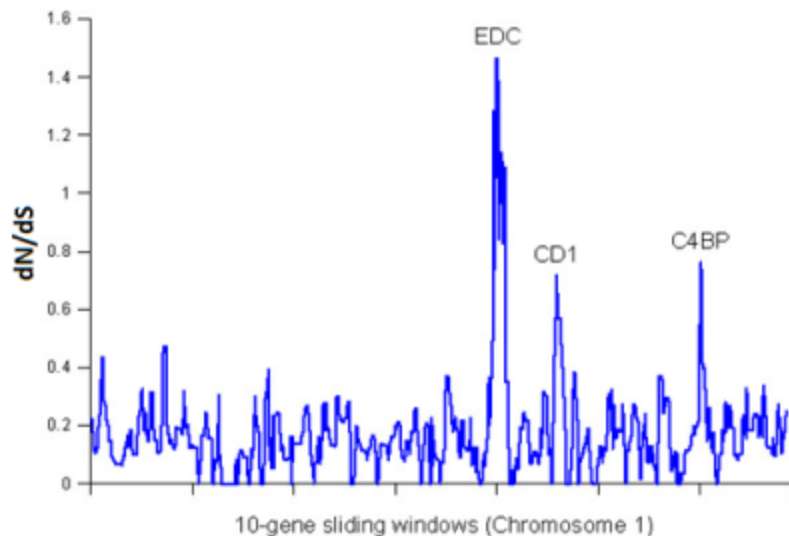
dN/dS in human-chimp

- Average dN/dS ~ 0.23

- 585 genes of 13,454 tested had dN/dS > 1
  - Genes involved in resistance to parasites, fertilization, etc.

- Shown in figure: Epidermal Differentiation Complex (EDC)

Try this!

- *Aspm* gene affects brain size in humans:
  - Human: CGT GGC GCA CTT GTA
  - Chimp: CGT CGT CCA CTT GTA

What does a dN/dS value *mean*?

- Real dN/dS for *Aspm* ~ 0.9 for whole gene
- Is this gene likely to be evolving *neutrally*?
  - Amino acid changes "don't matter"?

What does a dN/dS value *mean*?

- Real dN/dS for *Aspm* ~ 0.9 for whole gene
- Is this gene likely to be evolving *neutrally*?
  - Amino acid changes "don't matter"?

- **BOTH** constraint AND rapid evolution may operate within a gene (at different bases)
- This muddies a single "generalization" about a gene as a whole

What does a dN/dS value *mean*?

- If dN/dS = 1, then "cannot reject" neutrality
  - Realistically, unlikely any protein-coding gene would really be evolving totally neutrally

- If dN/dS < 1, then lots of "constraint"
  - *Most* amino acid changes disfavored

- If dN/dS > 1, then selection driving rapid change
  - *Multiple* amino acid changes favored

Need another test…

- dN/dS can be too conservative for finding adaptive amino-acid changes
  - Many false negatives

Image Credits, Unit 15-2

- EDC, © 2005 Republished with permission from Macmillan Publishers Ltd: Nature, The Chimpanzee Sequencing and Analysis Consortium (2005). "Initial sequence of the chimpanzee genome and comparison with the human genome." Nature 437, 69-87,