Bioinformatics Journal Club – 04/02/2003

DNA strand asymmetry

P Green et al. 2003. Nat. Genetics Apr 33(4): 514-7.

Presented by Ru-Fang Yeh

rufang@biostat.ucsf.edu
Division of Biostatistics, UCSF

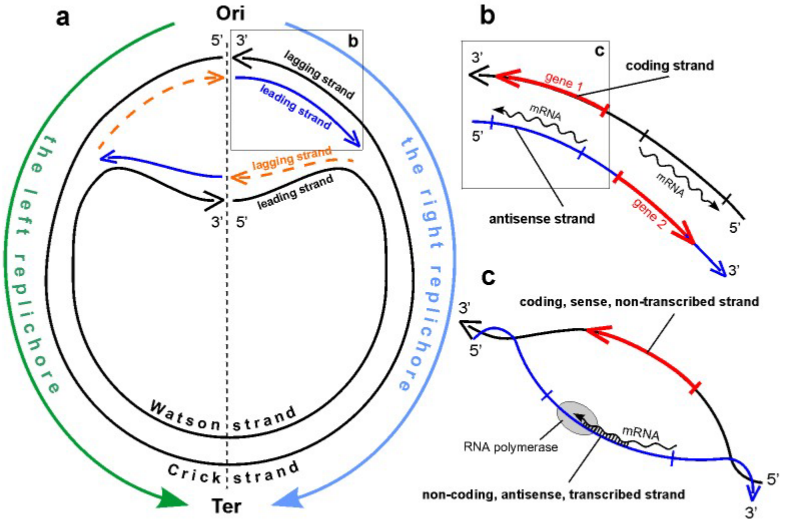
Chargaff's rules

- (first parity rule) for dsDNA:
  %A=%T, %C=%G
- second parity rule for ssDNA:
  %A≈%T, %C≈%G
  ➔ %purine(A/G) ≈ % pyrimidine(C/T)
  %K(G/T) ≈ %M(A/C)

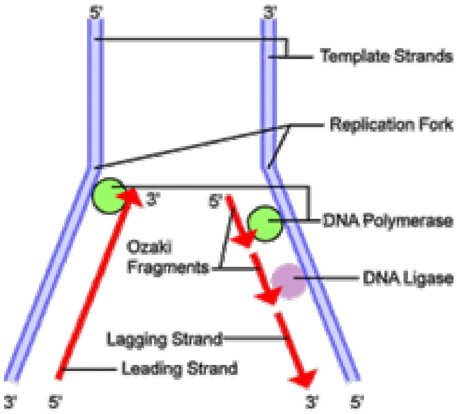
Strand asymmetry

- Mutation rate unequal on two strands
- substitution > complementary type
  eg. A ➔ G > T ➔ C

- Cause 1: DNA replication (bacteria, mitochondria)
- Cause 2: transcription-coupled repair & deamination (enterobacteria: E.coli etc.)

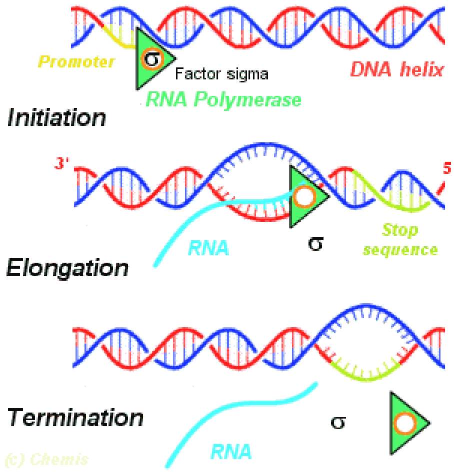


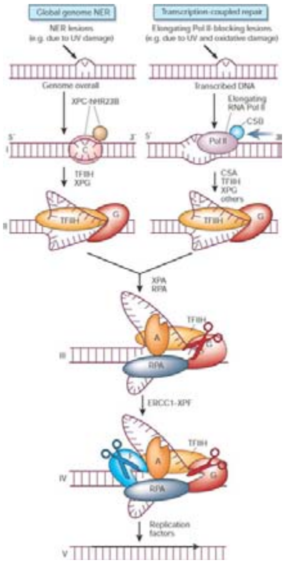
DNA Replication



Result: substitution asymmetry on leading vs lagging strand; across repl. origin

Transcription





Cyclobutane pyrimidine dimer: common DNA damage

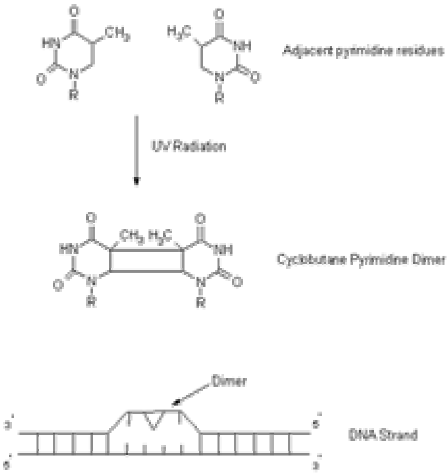
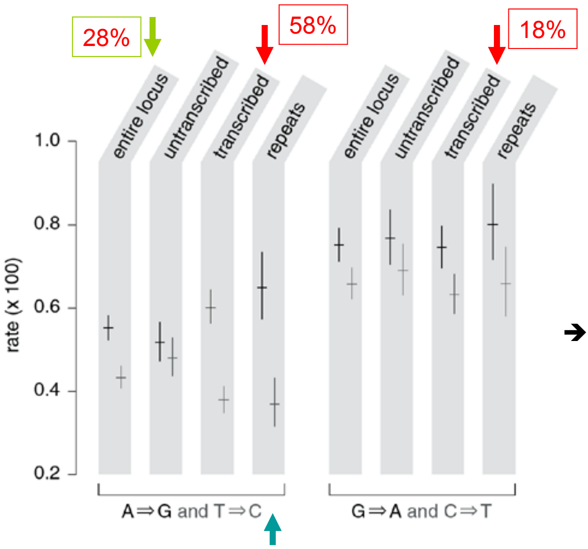


Fig. 2: human-chimpanzee



➔Asymmetrty in neutral mutation, NOT selection

Suppl. Fig. 1

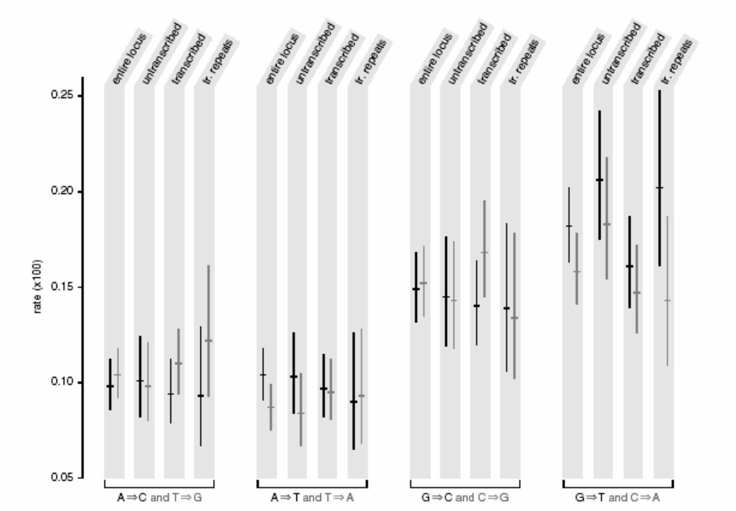


Fig. 3

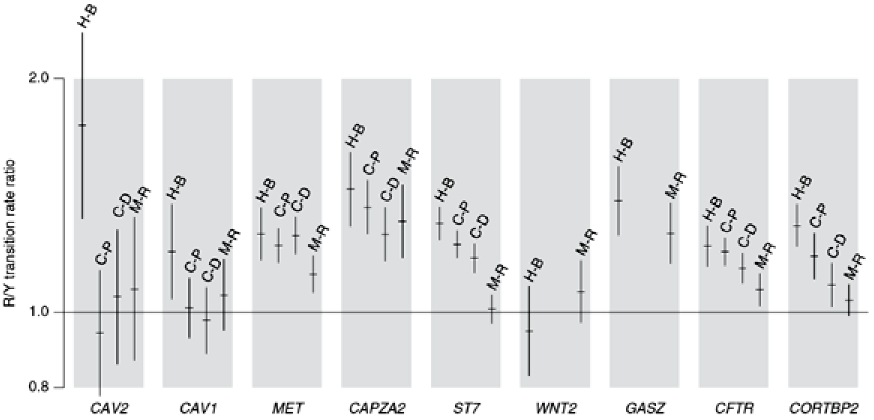
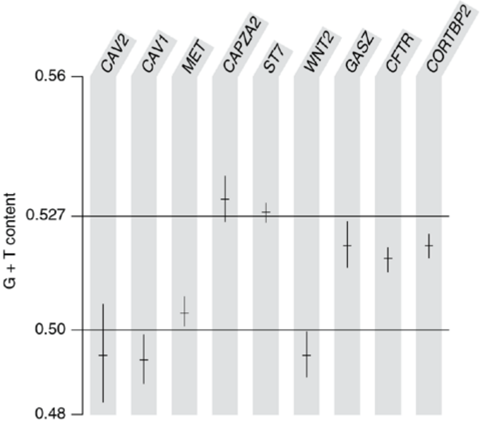
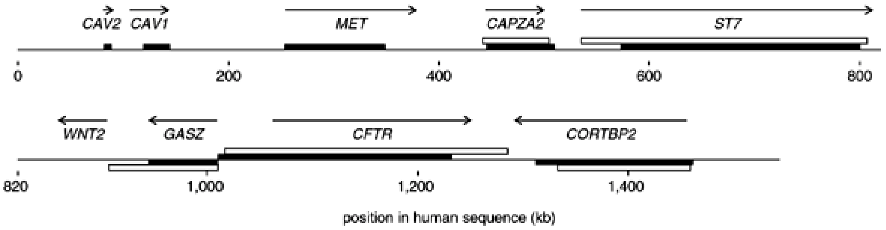


Fig 4



Since A→G / T→C >> G→A / C→T, over time G>>C & T >> A resulting an excess G+T: 52.7% equill.

Fig. 1: human-baboon



Maximal segment analysis using log odds score:

▬ • 55% purine vs random (50%) background; (0.138, -0.152)

▭ • 52% G+T vs random

• find high-scoring segments (score >12, type-I=.05) using BLAST-like algorithm

G+T excess in human Chr22

- 187 / 275 (68%) showed G+T excess
- > 10kb, 82% (159 / 195)
- > 20kb, 91% (127 / 139)
- Average G+T = 52.6%

Main finding

- No diff in overall substitution rate
- Excess purine transition & deficit of pyrimidine transition) on coding strand

  Note: Enterobacteria: excess of C→T coding-strand transition (explained by C deamination)
- Proposed mechanism: Transcription-coupled repair of uncorrected replication errors
- Potential use of closely related orgs for gene finding.