Research Staff

Yannis Almirantis, Research Director

Spyros Papageorgiou, Emeritus Scientist

Konstantinos Apostolou – Karampelis, Graduate Student

Dimitris Polychronopoulos, Collaborating Graduate Student

Research Associates and PhD or Master students in our group during the last years

Christoforos Nikolaou

Diamantis Sellis

Labrini Athanasopoulou

Alexandros Klimopoulos

Giannis Tsiagkas

Research Interests – general description

Probabilistic and statistical aspects in genome organization – Non-randomness at several length scales.

-

Deviations from randomness at the level of nucleotide n-tuplets. Patterns related to the functionality of genomic regions and to global genome structure.

-

Deviations from randomness at the “middle” length scale, expressed mainly through clustering of similar nucleotides. Distinction between protein-coding and non-coding functionalities.

-

Long range correlations and Zipf laws in the genome structure. Power-laws in the distribution of exons and of other genomic functional localizations. Entropic scaling in the study of genomic sequences as an indication of long range order and fractality.

-

DNA sequences seen as genomic text – Linguistic features in the genome: redundancy, multiple coding, polarity and asymmetry etc.

-

“Conservation laws” at the genome structure. The case of “Chargaff’s 2nd parity rule”. The use

of deviations from this law in the study of genomic dynamics and evolution.

-

Evolution at the genomic level. Formulation of minimal evolutionary scenarios compatible with the observed probabilistic features of genomes. Interpretation of the above mentioned probabilistic features either by selectionist or mutationist causality.

Pattern formation in biological systems – Self-organization and evolution.

-

Early development – Left-right asymmetries – Mechanisms of activation of Hox genes during limb development.

-

Reaction-diffusion systems – Spontaneous symmetry breaking and pattern-formation in systems with feedbacks and non-linear dynamics

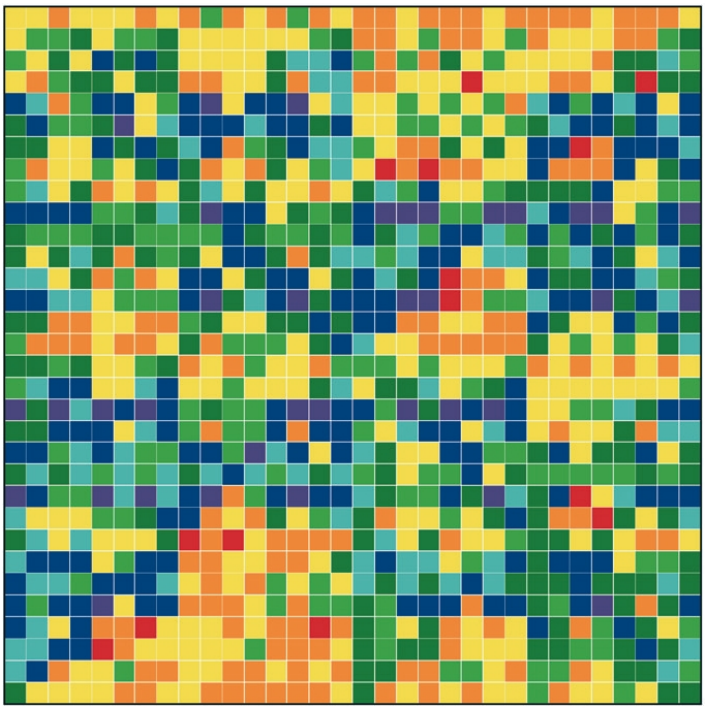
-

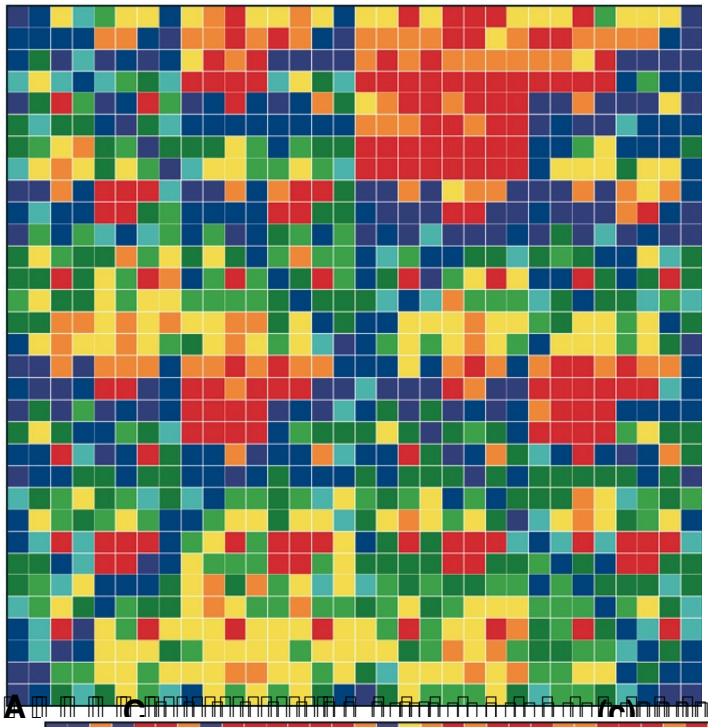
Prebiotic and early evolution as a complex self-organization procedure.

Findings in the field of Computational Genomics

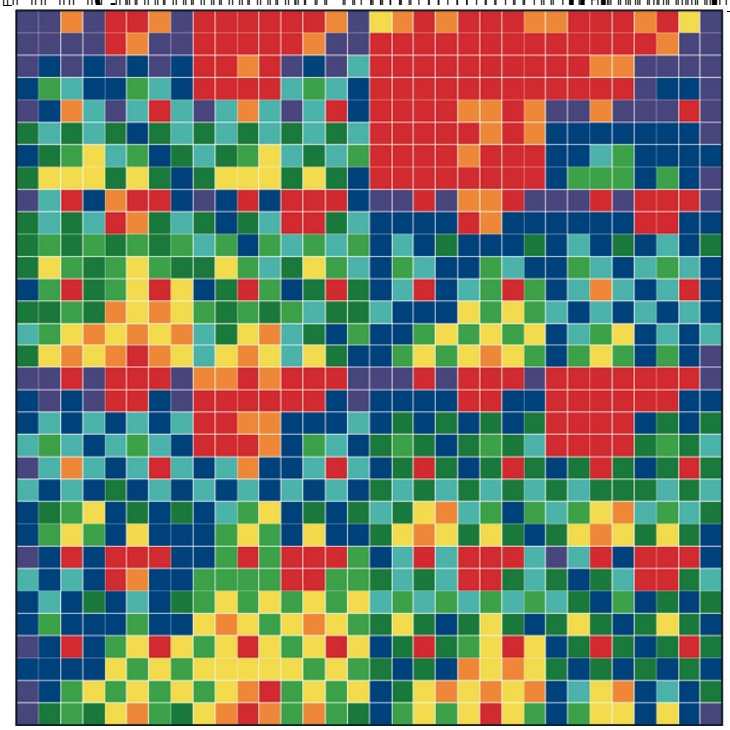
“Word” preference in the genomic text and genome evolution

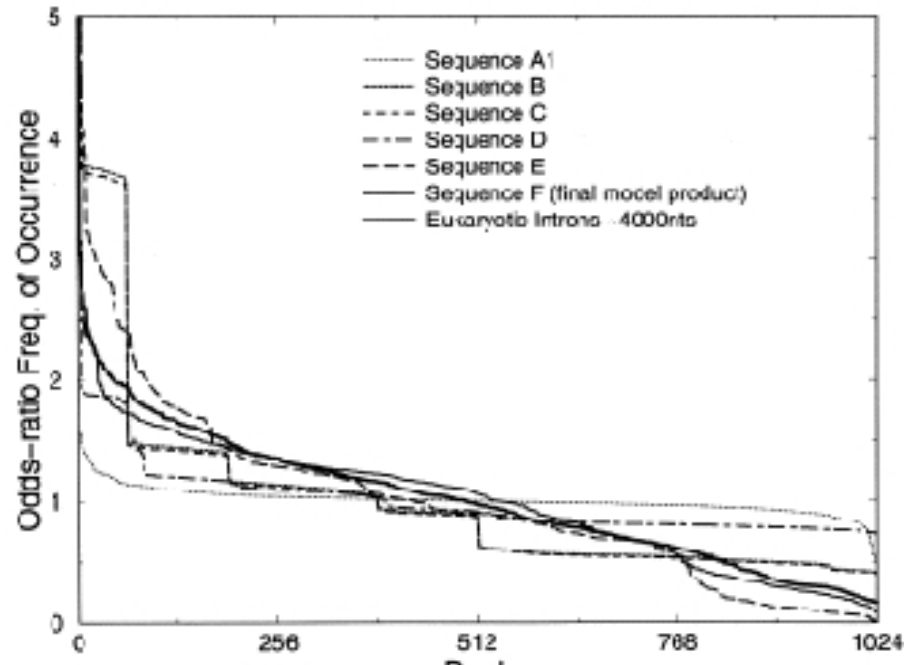
C (a) G C
(b) G

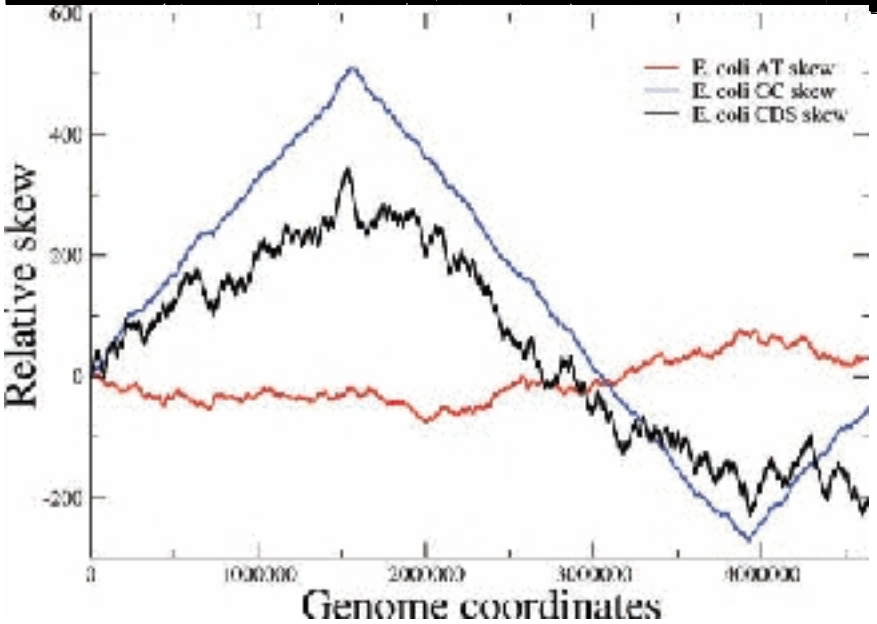




A C T T A (dG)




Sequence A1
Sequence B
Sequence C
Sequence D
Sequence E
Sequence F (final model produc)
Eukaryotic Introns 4000nts
Odds-ratio Freq. of Occurrence
Rank
0
1
2
3
4
5
256
512
768
1024


E. coli AT skew
E. coli GC skew
E. coli CDS skew
Relative skew
Genome coordinates
400
200
0
-200
1000000
2000000
3000000
4000000

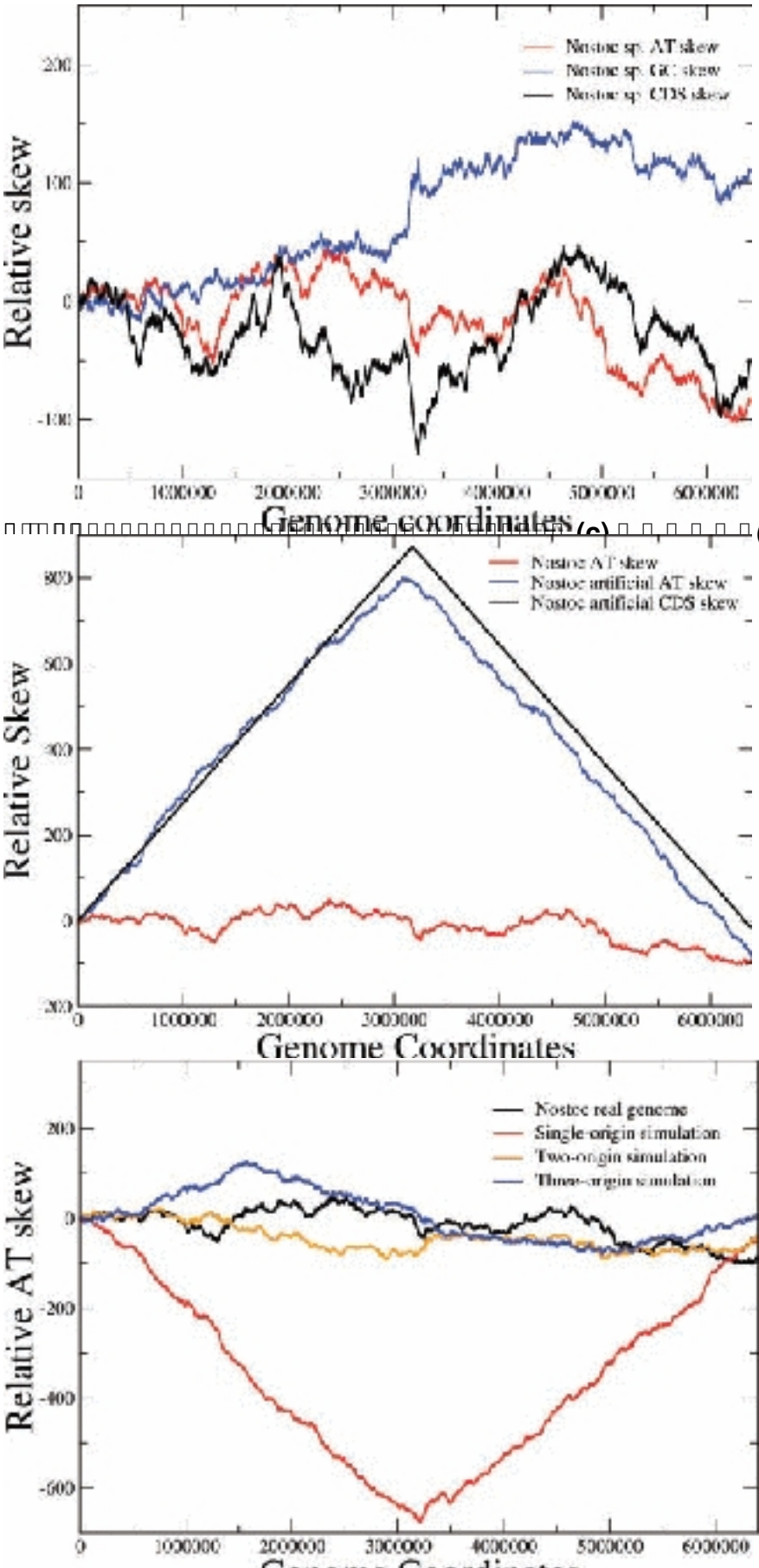
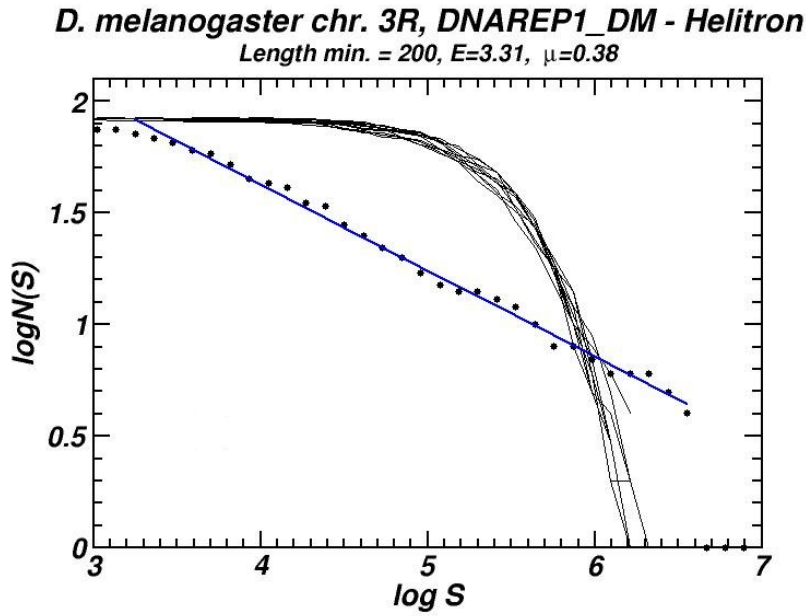
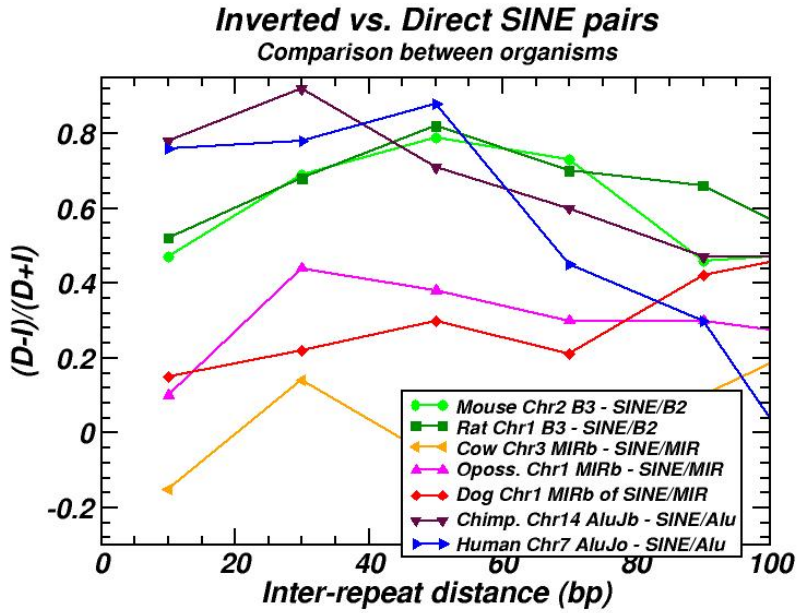
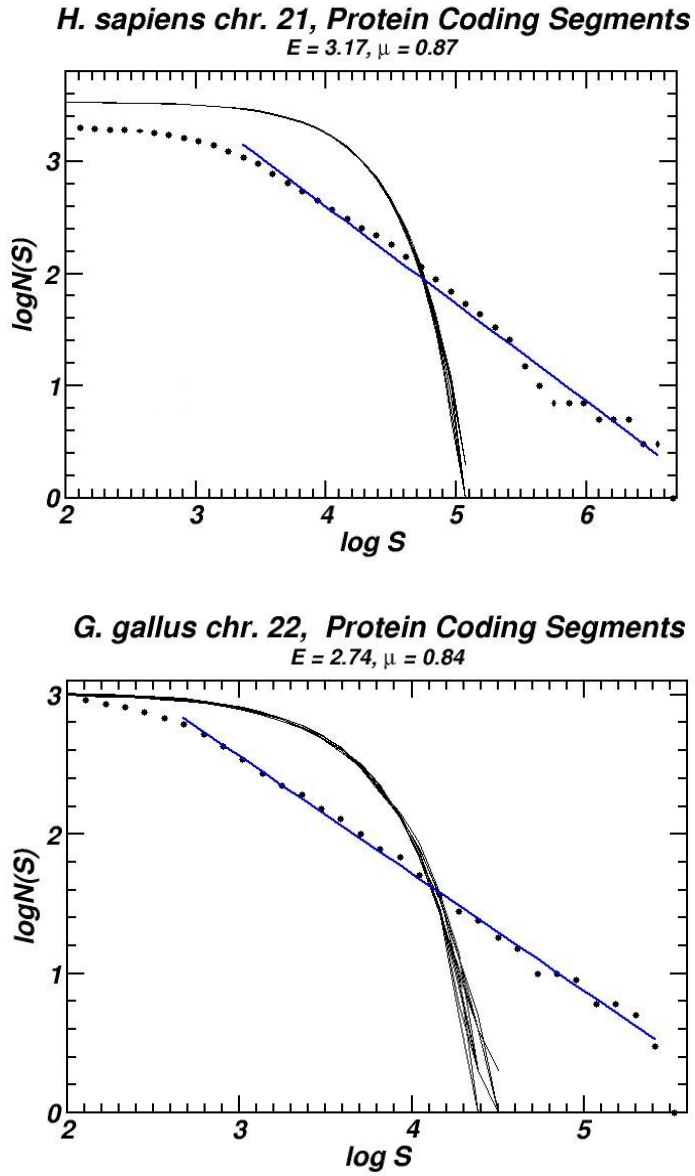


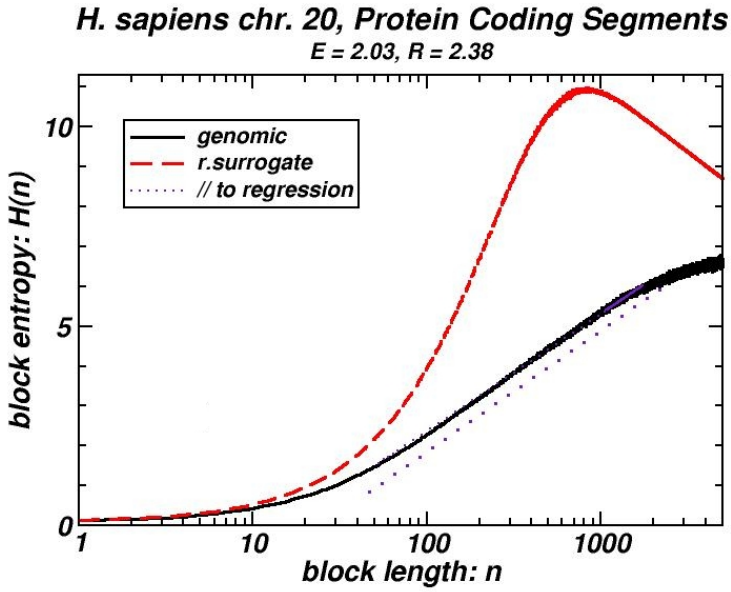
Figure 2. Example of bacterial genome AT skews. (a) Atypical eubacterium *E. coli* with a bilateral

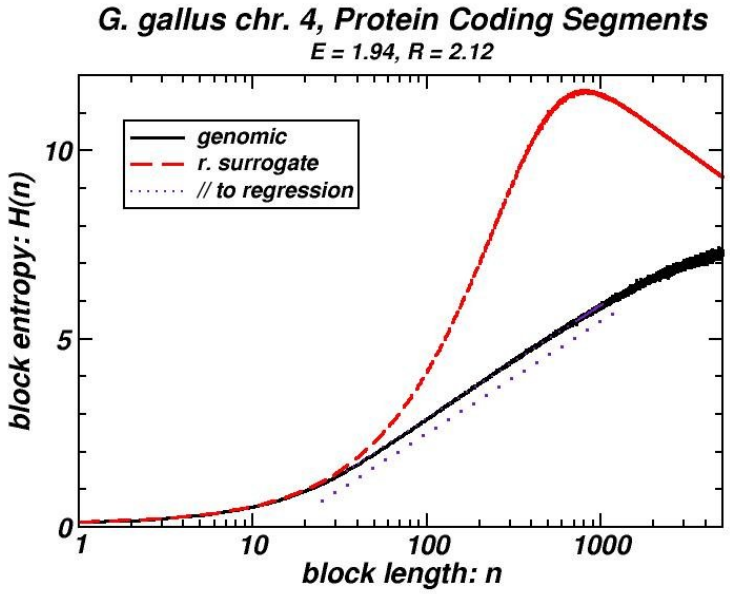

D. melanogaster chr. 3R, DNAREP1_DM - Helitron
Length min. = 200, E=3.31, μ=0.38
logN(S)
log S
0
0.5
1
1.5
2
3
4
5
6
7

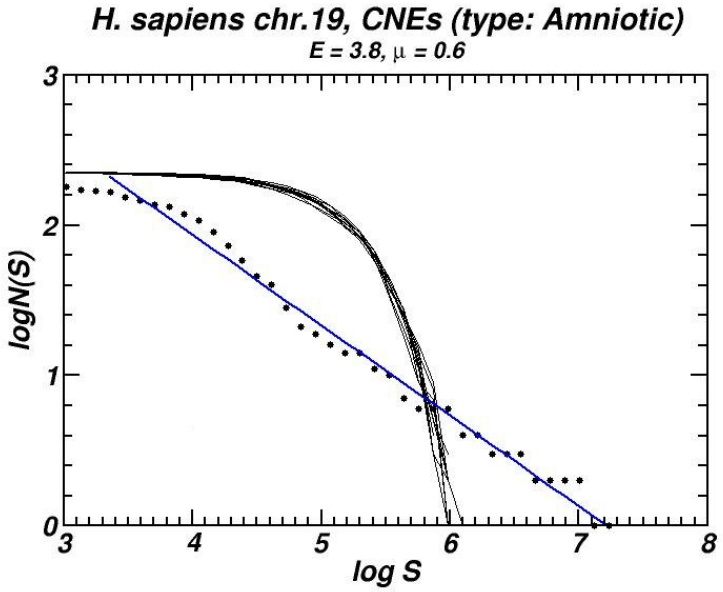

Inverted vs. Direct SINE pairs
Comparison between organisms
(D-I)/(D+I)
Inter-repeat distance (bp)
Mouse Chr2 B3 - SINE/B2
Rat Chr1 B3 - SINE/B2
Cow Chr3 MIRb - SINE/MIR
Oposs. Chr1 MIRb - SINE/MIR
Dog Chr1 MIRb of SINE/MIR
Chimp. Chr14 AluJb - SINE/Alu
Human Chr7 AluJo - SINE/Alu
-0.2
0
0.2
0.4
0.6
0.8
0
20
40
60
80
100

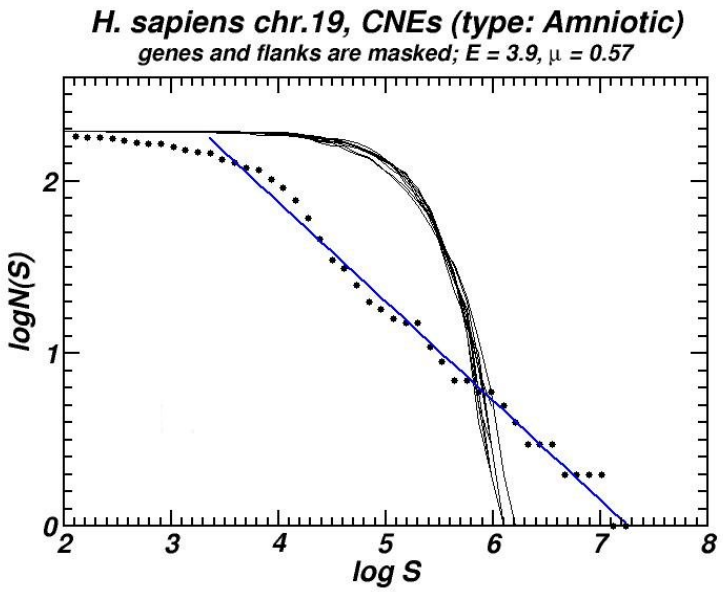

H. sapiens chr. 21, Protein Coding Segments
E = 3.17, μ = 0.87
logN(S)
log S
0
1
2
3
2
3
4
5
6
G. gallus chr. 22, Protein Coding Segments
E = 2.74, μ = 0.84
logN(S)
log S
0
1
2
3
2
3
4
5

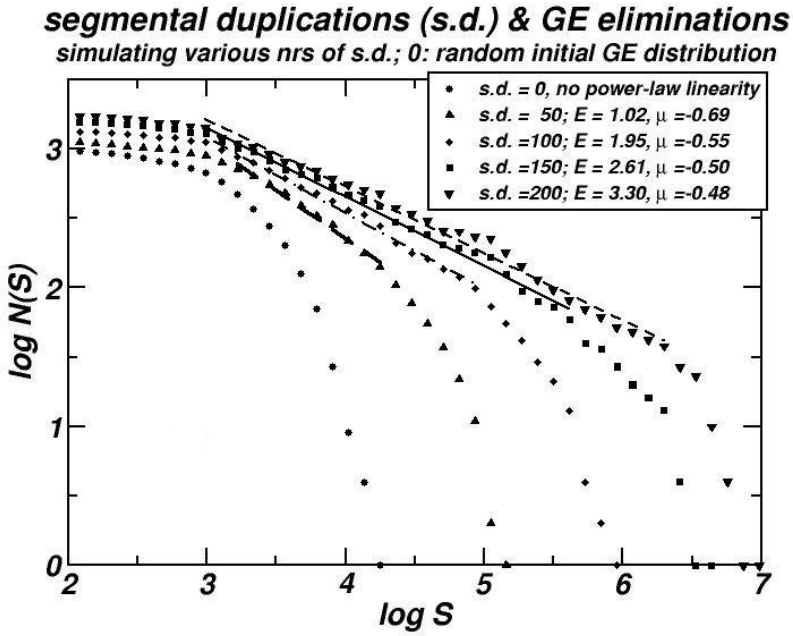
(a) **(b)**

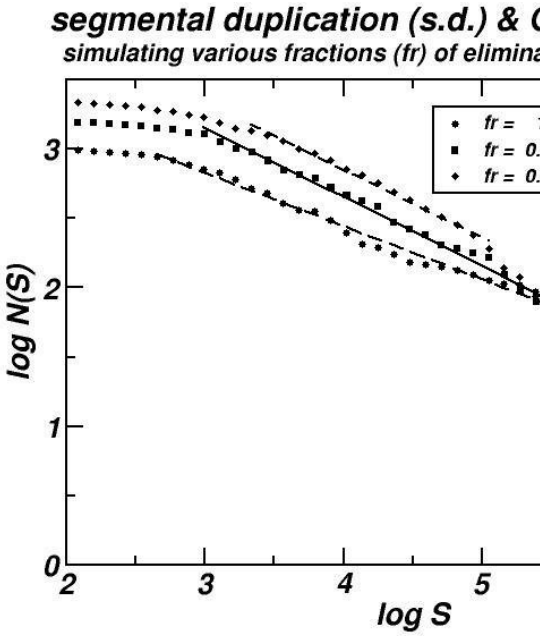

H. sapiens chr. 20, Protein Coding Segments
E = 2.03, R = 2.38
genomic
r.surrogate
// to regression
block entropy: H(n)
block length: n
0
5
10
1
10
100
1000


G. gallus chr. 4, Protein Coding Segments
E = 1.94, R = 2.12
genomic
r. surrogate
// to regression
block entropy: H(n)
block length: n
0
5
10
1
10
100
1000


H. sapiens chr.19, CNEs (type: Amniotic)
E = 3.8, μ = 0.6
logN(S)
log S
0
1
2
3
4
5
6
7
8


H. sapiens chr.19, CNEs (type: Amniotic)
genes and flanks are masked; E = 3.9, μ = 0.57
logN(S)
log S
0
1
2
3
4
5
6
7
8


segmental duplications (s.d.) & GE eliminations
simulating various nrs of s.d.; 0: random initial GE distribution
s.d. = 0, no power-law linearity
s.d. = 50; E = 1.02, μ =-0.69
s.d. =100; E = 1.95, μ =-0.55
s.d. =150; E = 2.61, μ =-0.50
s.d. =200; E = 3.30, μ =-0.48
log N(S)
log S
0
1
2
3
2
3
4
5
6
7


segmental duplication (s.d.) &
simulating various fractions (fr) of elimina
fr =
fr = 0.
fr = 0.
log N(S)
log S
0
1
2
3
2
3
4
5