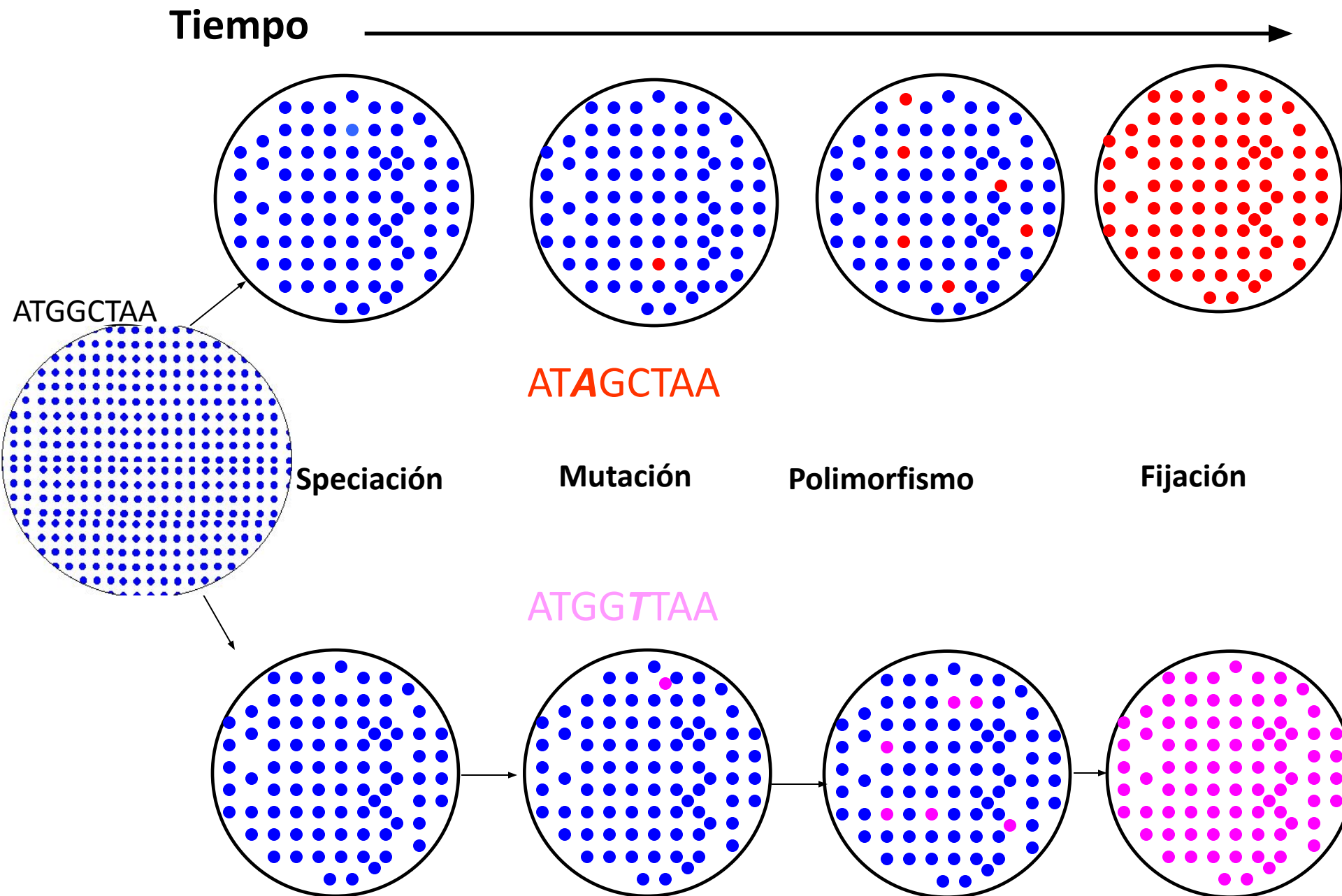
Teoría Neutralista de Kimura

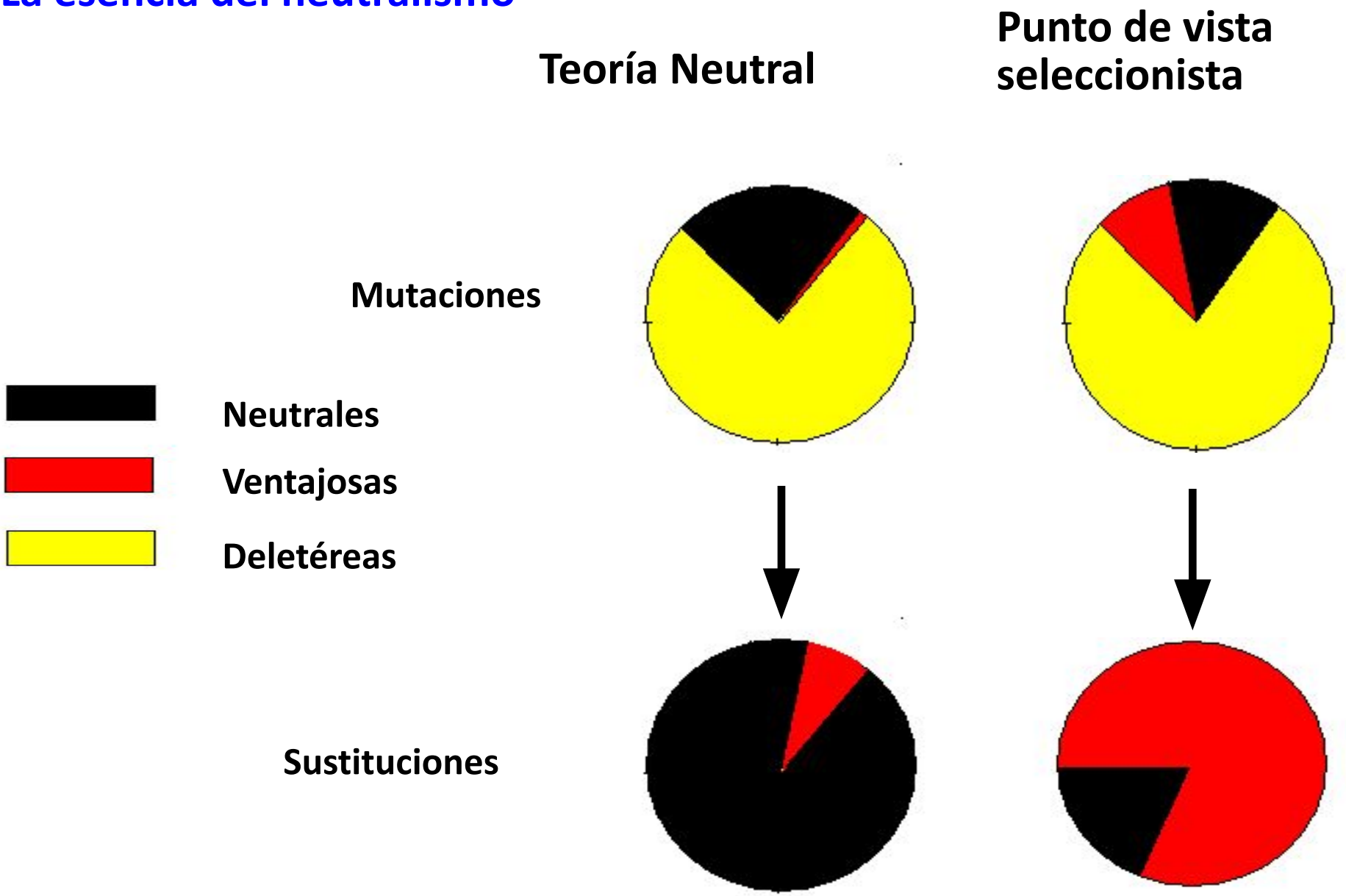
- La amplísima mayoría de las sustituciones son neutrales en relación al fitness
- Las mutaciones ventajosas son extremadamente raras
- La Deriva Genética domina la evolución a nivel molecular
- La tasa de evolución molecular es igual a la tasa de mutación neutral
- Las mutaciones deletéreas son eliminadas (la selección juega el rol de mantener el status quo)

MUTACIONES. POLIMORFISMOS Y SUSTITUCIONES



Comparación entre la Teoría Neutral y el Seleccionismo

La esencia del neutralismo



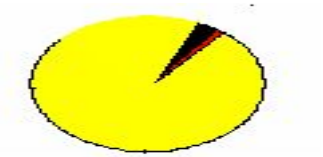
Algunas predicciones de la Teoría Neutral relacionadas a las tasas evolutivas

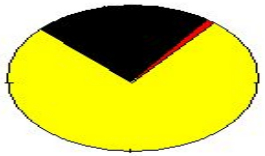
$$K = u \qquad u = m \bullet (1 - f)$$

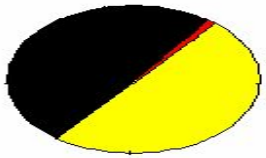
Neutrales
Ventajosas
Deletéreas

$$K = m(1 - f)$$

Mutaciones

1  ***f*=0.98** **(ej. histones)**

2  ***f*=0.75**

3  ***f*=0.50**

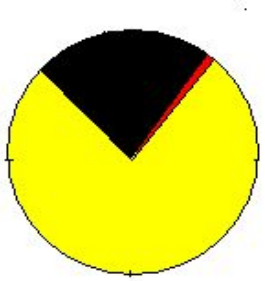
4  ***f*=0.20** **(ej. fibrinopeptidos)**

5  ***f*=0.0** **(ej. pseudogenes)**

Neutrales
Ventajosas
Deletéreas

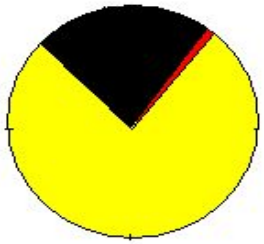
1

2

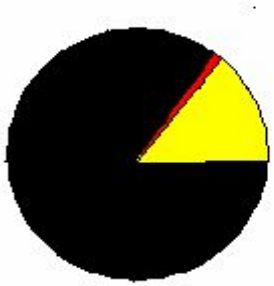


La secuencia # 2 debería evolucionar 25/2=12 veces más rápido que la secuencia 1

2

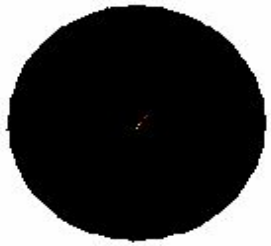


4



Mientrás que la secuencia 4 debería evolucionar 80/25=3.2 veces más rápido que la secuencia 2, y 40 veces más rápido que la secuencia 1

5



Finalmente una secuencia como la #5, se espera que tenga la tasa evolutiva más alta

Por lo que la comparación de las tasas evolutivas entre una secuencia dada con aquella completamente libre de restricciones funcionales permitiría estimar el valor de *f*

Mutation rates differ among regions of the mammalian genome

Kenneth H. Wolfe, Paul M. Sharp* & Wen-Hsiung Li†

Department of Genetics, Trinity College, Dublin 2, Ireland
† Center for Demographic and Population Genetics, University of Texas, PO Box 20334, Houston, Texas 77225, USA

In the traditional view of molecula mutation is uniform over the genom in the rate of nucleotide substitutio differential selective constraints[1,2]

Table 1 Rates of nucleotide substitution at silent sites and in functionless DNA for human versus Old World monkey sequences

Gene	$K_4 \pm$ s.e.m.	GC_4^*	L_4†
α_1-Antitrypsin	0.077 ± 0.022	73.6	184
Apolipoprotein A-I	0.061 ± 0.025	88.4	126
Apolipoprotein E	0.062 ± 0.021	89.1	175
Chorionic gonadotrophin-β	0.065 ± 0.027	81.3	99
Erythropoietin	0.110 ± 0.035	71.7	110
β-Globin	0.039 ± 0.028	68.2	54
Insulin	0.130 ± 0.054	82.3	62
Metallothionein I	0.125 ± 0.113	85.5	21
Metallothionein II	0.096 ± 0.072	87.0	23
β-Myosin heavy chain	0.179 ± 0.056	86.5	97
Pepsinogen A	0.108 ± 0.026	73.7	194
Transforming growth factor-β	0.064 ± 0.019	83.5	203
Triose phosphate isomerase	0.112 ± 0.033	59.8	128
Total for silent sites	0.089 ± 0.009	78.7	1,474
η-Globin pseudogene locus	0.074 ± 0.006	42.9	2,071
η-δ Globin intergenic DNA	0.076 ± 0.006	38.7	2,771
Total for functionless DNA	0.075 ± 0.004	40.5	4,842

K_4 (the corrected number of substitutions per fourfold degenerate site) was calculated as described in Fig. 1. References to the original DNA sequence data are available on request.

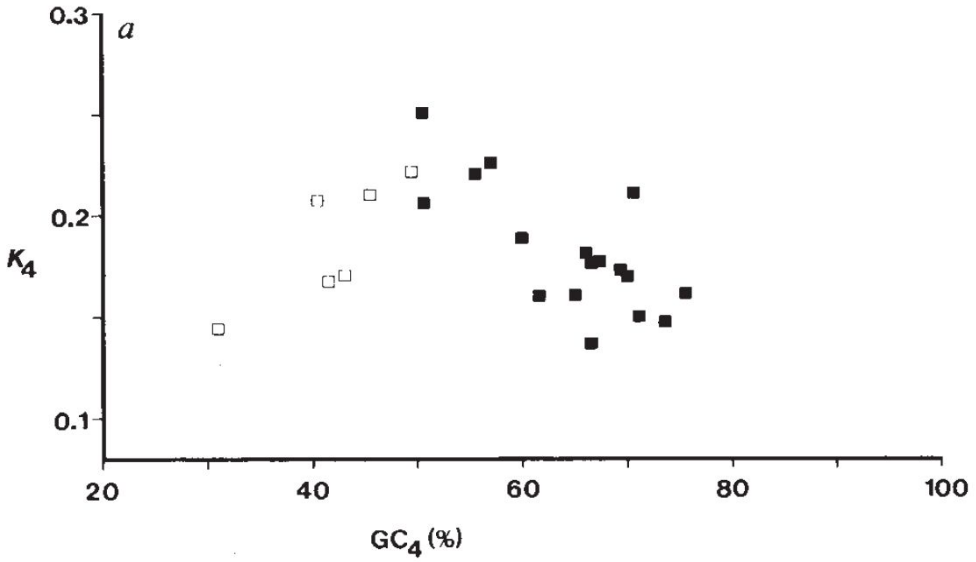
Mutation rates differ among regions of the mammalian genome

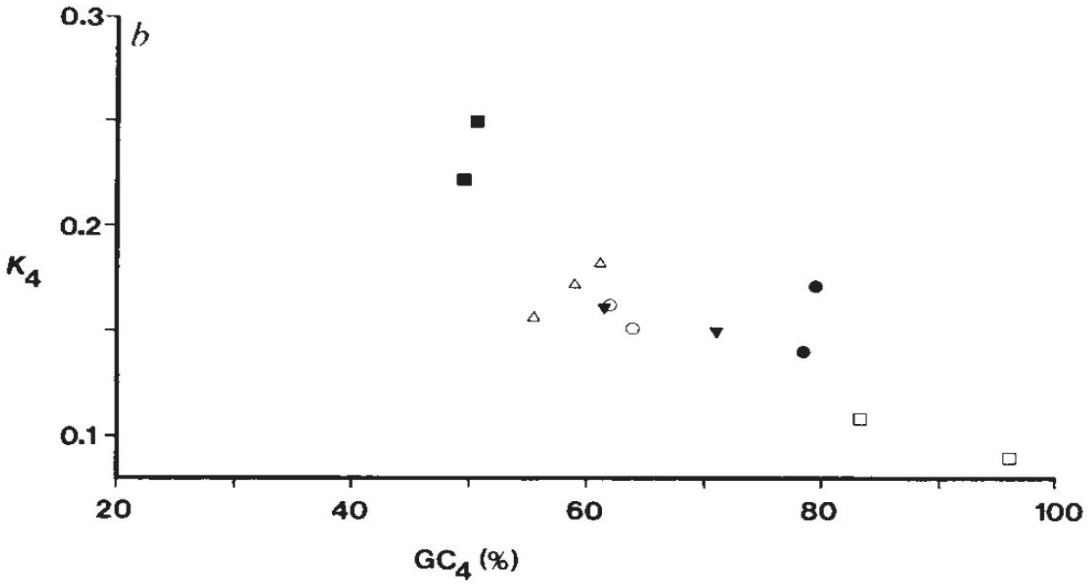
Kenneth H. W

Department of Ge
† Center for Demo
Texas, PO Box 20

In the traditiona
mutation is unifo
in the rate of nuc
differential selec

Table 1 Rates
less DNA





Genes duplicados

The variance of GC_4 in the 23 large rodent genes (Fig. 1*a*) is more than ten times that expected by chance (under a binomial distribution with mean 58.6%), suggesting that the factors determining G+C content are not uniform for all genes. Of course, one might argue that the GC_4 value in any gene is at a particular optimum and that the observed relationship between K_4 and GC_4 is a consequence of selection maintaining that GC_4 value in the face of a mutation rate that is uniform across the genome. But this would require strong selective constraints, whereas we have shown here that silent sites seem to be effectively neutral. Furthermore, the direction of selection on G+C-rich genes would have to be opposite to that on A+T-rich genes. Thus our finding that the substitution rate and the base composition of silent sites vary together in a systematic way is most simply explained by supposing that the pattern of mutation is different for different genes. Most germline mutations are thought to arise from misincorporation errors made by the DNA replication apparatus[14,15]. It has been demonstrated that different genes replicate at different stages of the cell cycle in differentiated cells[16], and this is presumably also true in the germline. The number and type of replication errors are likely to vary during the cell cycle if the chemical environment in the nucleus changes. In fact, the abundances (both relative and absolute) of free dNTPs in the nucleus change with time[17], as do the activities of the DNA polymerase enzymes and their accessory proteins[18]. We have examined a theoretical model (to be detailed elsewhere) of the relationship between the mutation rate and the G+C

ably also to high-resolution Giemsa chromosomal bands[16]. We therefore propose that isochores arise as a result of the synchronous replication of megabase stretches of DNA under varying dNTP pool conditions. Although our model based on variation in the dNTP precursor pools can provide a simple explanation for the observed variation of mutation rates and patterns around the genome, it is of course not the only possible explanation. For example, Filipski[22] has proposed that isochores are formed as a consequence of the repair of DNA in different types of chromatin by different DNA polymerase enzymes, and there is now some experimental evidence for between-gene differences in efficiency of DNA repair[23]. Our observations could also be explained if DNA is replicated by several distinct DNA polymerase holoenzymes with different error propensities. But recent data indicate that there is only one replicative polymerase complex for mammalian nuclear DNA[18,24].

J Mol Evol (1982) 18:360–369

Journal of
Molecular Evolution

Patterns of Nucleotide Substitution in Pseudogenes and Functional Genes

Takashi Gojobori, Wen-Hsiung Li, and Dan Graur

Center for Demographic and Population Genetics, University of Texas Health Science Center, Houston, Texas 77030, USA

Summary. The pattern of point mutations is inferred from nucleotide substitutions in pseudogenes. The pattern obtained suggests that transition mutations occur somewhat more frequently than transversion mutations and that mutations result more often in A or T than in G or C. Our results are discussed with respect to the predictions from Topal and Fresco's model for the molecular basis of point (substitution) mutations (Nature 263:285–289, 1976). The pattern of nucleotide substitution at the first and second positions of codons in functional genes is quite similar to that in pseudogenes, but the relative frequency of the transition C→T in the sense strand is drastically reduced and those of the transversions C→G and G→C are doubled. The differences between the two patterns can be explained by the observation that in the protein evolution amino acid substitutions occur mainly between amino acids with similar biochemical properties (Grantham, Science 185:862–864, 1974). Our results for the patterns of nucleotide substitutions in pseudogenes and in functional genes lead to the prediction that both the coding and non-coding regions of protein coding genes should have high frequencies of A and T. Available data show that the non-coding regions are indeed high in A and T but the coding regions are low in T, though high in A.

J Mol Evol (1982) 18:360–369

Journal of
Molecular Evolution

Patterns of Nucleotide Substitution in Pseudogenes and Functional Genes

Takashi Gojobori, Wen-Hsiung Li, and Dan Graur

Center for Demographic and Population Genetics, University of Texas Health Science Center, Houston, Texas 77030, USA

often only in certain directions. Here we propose to infer this pattern from DNA sequences for pseudogenes. As pseudogenes are apparently subject to no functional constraint, all mutations in them would be selectively neutral and would become fixed in the population with equal probability. Thus the pattern of nucleotide substitutions in pseudogenes would reflect the pattern of spontaneous substitution mutations.

J Mol Evol (1982) 18:360–369

Journal of
Molecular Evolution

Patterns of Nucleotide Substitution in Pseudogenes and Functional Genes

Takashi Gojobori, Wen-Hsiung Li, and Dan Graur

Center for Demographic and Population Genetics, University of Texas Health Science Center, Houston, Texas 77030, USA

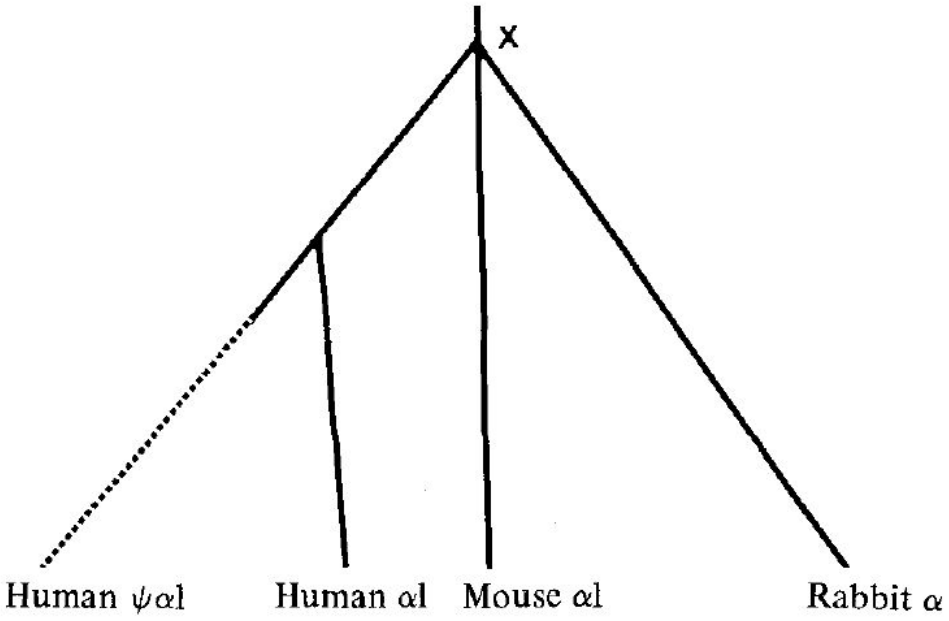


Fig. 1. Plausible phylogenetic tree for human ψαl, human αl, mouse αl, and rabbit α

Table 1. Proportions of base substitutions in pseudogenes and in functional genes

Comparison	A→T	A→C	A→G	T→A	T→C	T→G	C→A	C→T	C→G	G→A	G→T	G→C	Total
Pseudogenes: all positions of codons													
Human ψα1	1/68	1/68	7/68	3/65	2/65	1/65	13/136	25/136	5/136	13/113	2/113	6/113	79/382 = 0.21
Mouse ψα3	1/70	2/70	5/70	3/79	1/79	5/79	1/111	9/111	1/111	7/88	3/88	1/88	39/348 = 0.11
Rabbit ψβ2	2/88	4/88	5/88	3/104	6/104	3/104	6/101	10/101	3/101	12/132	4/132	5/132	63/425 = 0.15
Goat $\psi\beta^X$	1/45	1/45	3/45	1/43	0/43	0/43	2/44	7/44	3/44	9/69	3/69	3/69	33/201 = 0.16
Mouse ψβh3	3/28	1/28	1/28	0/26	0/26	0/26	2/31	3/31	0/31	4/33	3/33	0/33	17/118 = 0.14
Human ψV	2	3	3	0	7	2	3	9	3	7	6	3	48

Table 2. Relative substitution frequencies (%) in pseudogenes and functional genes

	Pseudogenes										Functional genes				
	Human ψα1					Human ψV_κ					α genes (Hα1, Mα1, Rα)				
	A	T	C	G	[59.5]	A	T	C	G	[54.1]	A	T	C	G	[44.6]
A	---	2.0	2.0	14.2	18.2	---	4.5	6.8	6.8	18.1	---	5.0	0.0	12.6	17.6
T	6.4	---	4.2	2.1	12.7	0.0	---	15.4	4.4	19.8	3.2	---	6.5	0.0	9.7
C	13.2	25.3	---	5.1	43.6	5.9	17.6	---	5.9	29.4	9.6	4.8	---	14.3	28.7
G	15.8	2.4	7.3	---	25.5	14.3	12.3	6.1	---	32.7	20.7	2.6	20.7	---	44.0
	35.4	29.7	13.5	21.4		20.2	34.4	28.3	17.1		33.5	12.4	27.2	26.9	

J Mol Evol (1982) 18:360–369

Journal of
Molecular Evolution

Patterns of Nucleotide Substitution in Pseudogenes and Functional Genes

Takashi Gojobori, Wen-Hsiung Li, and Dan Graur

Center for Demographic and Population Genetics, University of Texas Health Science Center, Houston, Texas 77030, USA

Average (8 pseudogenes)*

	A	T	C	G	[55.6] (52.0)
A	---	4.7±1.9 (5.3±2.3)	5.2±0.8 (5.7±0.8)	11.4±1.6 (12.1±1.8)	21.3 (23.1)
T	4.5±1.0 (4.7±1.1)	---	6.2±1.8 (6.7±1.9)	4.6±1.8 (5.1±2.1)	15.3 (16.5)
C	8.3±1.4 (9.3±1.7)	22.0±1.8 (18.0±1.5)	---	4.7±1.0 (4.8±1.1)	35.0 (32.1)
G	16.0±1.1 (15.2±1.4)	7.0±1.5 (7.7±1.8)	5.5±0.8 (5.4±0.8)	---	28.5 (28.3)
	28.8 (29.2)	33.7 (31.0)	16.9 (17.8)	20.7 (22.0)	

Average (3 functional genes)

	A	T	C	G	[45.5]
A	---	4.2±2.0	6.3±2.8	11.5±1.4	22.0
T	5.0±1.2	---	3.7±1.6	1.7±1.4	10.4
C	8.1±0.6	9.4±3.3	---	13.0±1.5	30.5
G	20.9±3.1	4.8±0.8	11.5±3.8	---	37.2
	34.0	18.4	21.5	26.2	

* The values in parentheses are obtained by excluding the nucleotide sites were the CG dinucleotide appeared to have occurred in the ancestral sequences of these pseudogenes (see text)

Synonymous Codon Bias Is Not Caused by Mutation Bias in G+C-Rich Genes in Humans

Nick G. C. Smith and Adam Eyre-Walker

Centre for the Study of Evolution and School of Biological Sciences, University of Sussex, Brighton, England

It is has been suggested that synonymous codon bias is a consequence of mutation bias in mammals. We tested this hypothesis in humans using single-nucleotide polymorphism data. We found a pattern of polymorphism which was inconsistent with the mutation bias hypothesis in G+C-rich genes. However, the data were consistent with the action of natural selection or biased gene conversion. Similar patterns of polymorphism were also observed in noncoding DNA, suggesting that natural selection or biased gene conversion may affect large tracts of the human genome.

Introduction

It is well established that selection acts on synonymous codon use in many groups of organisms, including bacteria, gating in a sample is expected to be equal to the number of GC→AT mutations if mutation bias is the sole cause

Table 1
The Numbers of GC→AT and AT→GC Synonymous Mutations Segregating in Human Genes

GC_3	GC→AT	AT→GC	*P*
0.20–0.30	2	0	NS
0.30–0.40	9	7	NS
0.40–0.50	9	12	NS
0.50–0.60	8	7	NS
0.60–0.70	19	4	0.003
0.70–0.80	18	4	0.005
0.80–0.90	23	3	0.0001
Total.	88	37	6×10^{-6}

NOTE.—The data are divided according to the GC_3 (G+C content at the third codon position) of the exon containing the single-nucleotide polymorphisms, and the *P* value is for the test of $M_{GC\rightarrow AT} = M_{AT\rightarrow GC}$ obtained from a binomial distribution $B[M_{GC\rightarrow AT} + M_{AT\rightarrow GC}, 0.5]$.

ELSEVIER

Gene 300 (2002) 161–168

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Isochores, GC_3 and mutation biases in the human genome

Fernando Alvarez-Valin[a], Guille

[a]Seccion Biomatematica, Facultad
[b]Laboratorio di Evoluzione Molecolare, Stazione Zoolo

Received 12 June 2002; received in revised
Received

Table 2
Mutation databases used in this study

	GC_3	Number of
1. GC_3-poor genes ($GC_3 < 0.45$)		
1.1. Haemophilia B factor (factor IX)	0.338	1474
http://www.umds.ac.uk/molgen/haemBdatabase.htm		
1.2. Haemophilia A factor (factor VIII)	0.388	428
http://europium.csc.mrc.ac.uk/usr/WWW/WebPages/main.dir/main.htm		
1.3. Ataxia telangiectasia	0.322	126
http://www.vmresearch.org/atm.htm		
1.4. Cystic fibrosis[a]	0.3975	509
http://www.genet.sickkids.on.ca/cftr/		
1.5. HPRT (Lesch–Nyhan syndrome)	0.3791	168
http://www.ibiblio.org/dnam/mainpage.html		
2. Genes with intermediate GC_3 content ($0.45 < GC_3 < 0.6$)		
2.1. Phenylalanine hydroxylase locus (Phenylketonuria)[a]	0.519	262
http://ww2.mcgill.ca/pahdb/		
2.2. PHEX X-linked hypophosphatemia	0.457	64
http://data.mch.mcgill.ca/phexdb/		
2.3. PAX6 Developmental eye anomalies	0.5174	80
http://www.hgu.mrc.ac.uk/Softdata/PAX6/		
3. GC_3-rich genes ($0.6 \leq GC_3 < 0.75$)		
3.1. Androgen receptor[b]	0.64	314
http://ww2.mcgill.ca/androgendb/		
3.2. P53 gene[c]	0.61	141
http://www.iarc.fr/p53/		
3.3. Wilson disease	0.6057	125
http://www.medgen.med.ualberta.ca/database.html		
4. Very GC_3-rich genes ($GC_3 > 0.75$)		
4.1. Glucose-6-phosphate dehydrogenase (Favism)	0.84	118
http://rialto.com/favism/mutat.htm		
4.2. L1CAM, L1 cell adhesion molecule	0.77	65
Van Camp et al. (1996)		
4.3. Haemophilia factor VII	0.7982	148
http://europium.csc.mrc.ac.uk/usr/WWW/WebPages/FVII/database.dir/index.htm		
4.4. LDLR locus (familial hypercholesterolaemia)	0.76	359
http://www.ucl.ac.uk/fh/		
Total		4381

ELSEVIER

Gene 300 (2002) 161–168

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

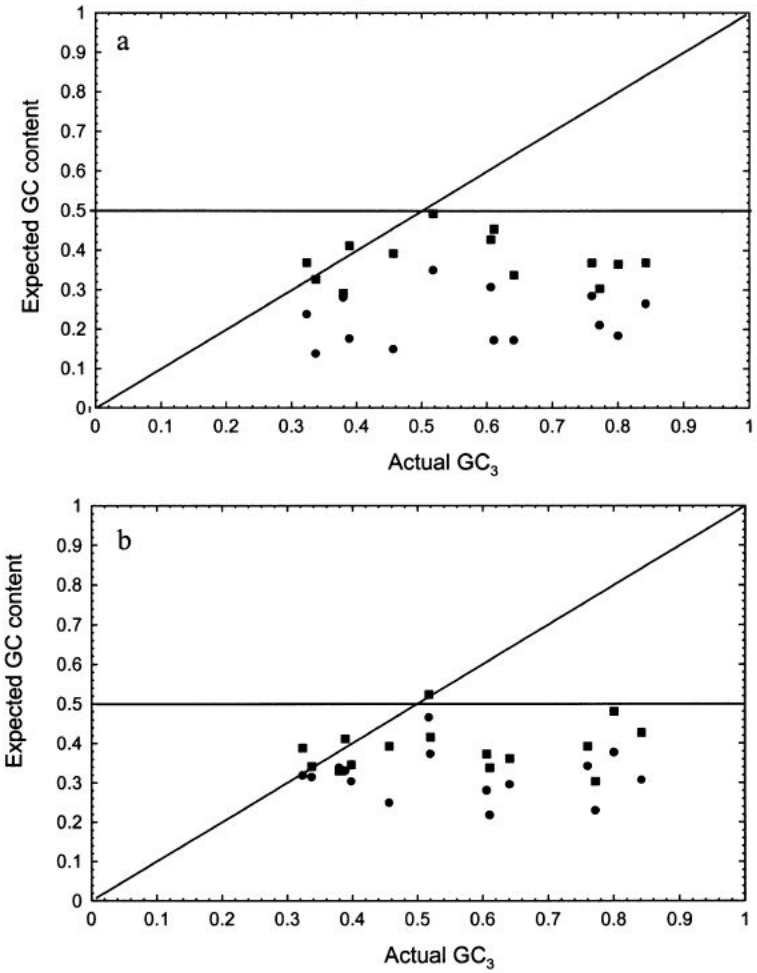
www.elsevier.com/locate/gene

Isochores, GC_3 and mutation biases in the human genome

Fernando Alvarez-Valin[a], Guillermo Lamolle[a], Giorgio Bernardi[b,*]

[a]*Seccion Biomatematica, Facultad de Ciencias, Montevideo, Uruguay*
[b]*Laboratorio di Evoluzione Molecolare, Stazione Zoologica Anton Dohrn, Villa Comunale, I-80121 Naples, Italy*

Received 12 June 2002; received in revised form 4 July 2002; accepted 18 September 2002
Received by T. Gojobori



Vanishing GC-Rich Isochores in Mammalian Genomes

Laurent Duret,*,[1] Marie Semon,* Gwenaël Piganeau,[†] Dominique Mouchiroud* and Nicolas Galtier[‡]

**Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558 Université Claude Bernard Lyon 1, 69622 Villeurbanne Cedex, France, [†]Centre for the Study of Evolution, School of Biological Sciences, Falmer, Brighton BN1 9QG, United Kingdom and [‡]Laboratoire Génome, Populations, Interactions, UMR CNRS 5000 Université Montpellier 2, 34095 Montpellier Cedex 5, France*

Manuscript received March 12, 2002
Accepted for publication September 9, 2002

ABSTRACT

To understand the origin and evolution of isochores—the peculiar spatial distribution of GC content within mammalian genomes—we analyzed the synonymous substitution pattern in coding sequences from closely related species in different mammalian orders. In primate and cetartiodactyls, GC-rich genes are undergoing a large excess of GC → AT substitutions over AT → GC substitutions: GC-rich isochores are slowly disappearing from the genome of these two mammalian orders. In rodents, our analyses suggest both a decrease in GC content of GC-rich isochores and an increase in GC-poor isochores, but more data will be necessary to assess the significance of this pattern. These observations question the conclusions of previous works that assumed that base composition was at equilibrium. Analysis of allele frequency in human polymorphism data, however, confirmed that in the GC-rich parts of the genome, GC alleles have a higher probability of fixation than AT alleles. This fixation bias appears not strong enough to overcome

Pattern of synonymous substitutions (AT ↔ GC) in mammalian genes of different GC content

Order	GC3 class (%)	All codons			Quartet codons		No CpG:	GC3q[d] (%)	GC3eq[e] (%)
		GC → AT[a]	AT → GC[a]	Ratio[b]	GC → AT[a]	AT → GC[a]	GC → AT[c]		
Rodentia	<57	934	1378	0.7	374	585	327	46	58
	57–75	2143	1923	1.1	910	904	806	64	61
	>75	431	249	1.7**	182	128	157	78	69
Cetartiodactyla	<57	155	161	1	65	65	53	44	46
	57–75	185	104	1.8**	97	45	85	64	48
	>75	162	50	3.2**	88	20	64	83	56
Primates	<57	40	25	1.6*	18	11	13	41	34
	57–75	53	32	1.7*	31	23	24	63	54
	>75	37	10	3.7**	18	4	13	80	43

Data sets were split into three groups of genes of low, medium, and high GC3 content.

[a] Total number of synonymous substitutions counted in the branches leading to the two ingroups (see Table 1 for details).

[b] Ratio of GC → AT over AT → GC substitutions. Significance was assessed by a binomial test (probability of observing that ratio or higher assuming an equal expected number of the two kinds of changes). *$P < 0.05$; **$P < 0.01$.

[c] Number of GC → AT substitutions at the third position of quartet codons (*i.e.*, fourfold degenerate codons), excluding all positions corresponding to a CpG dinucleotide in the ancestral sequence.

[d] GC3q, average GC content at the third position of quartet codons.

[e] GC3eq, GC content expected at equilibrium at the third position of quartet codons (see text). Note that given the low number of substitutions analyzed in primate GC-rich genes, this estimate might not be very accurate for that subset.

Vanishing GC-Rich Isochores in Mammalian Genomes

Laurent Duret,*,[1] Marie Semon,* Gwenaël Piganeau,[†] Dominique Mouchiroud* and Nicolas Galtier[‡]

**Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558 Université Claude Bernard Lyon 1, 69622 Villeurbanne Cedex, France, [†]Centre for the Study of Evolution, School of Biological Sciences, Falmer, Brighton BN1 9QG, United Kingdom and [‡]Laboratoire Génome, Populations, Interactions, UMR CNRS 5000 Université Montpellier 2, 34095 Montpellier Cedex 5, France*

Manuscript received March 12, 2002
Accepted for publication September 9, 2002

TABLE 3

Pattern of synonymous substitutions (AT ↔ GC) in primate genes located in different isochore contexts

GC class[a] (%)	GC → AT[b]	AT → GC[b]	Ratio[c]
<43	60	39	1.5
43–49	40	18	2.2
>49	30	10	3.0

[a] Genes were split into three groups according to the GC content of the GenBank genomic sequence containing them (average length of the genomic sequence = 163 ± 50 kb). The limits of GC content correspond to the 33% lowest and highest GC content in the entire data set of 1892 complete human genes (see Figure 2 legend).

[b] Total number of synonymous substitutions counted in the branches leading to the two ingroups.

[c] Ratio of GC → AT over AT → GC substitutions.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 31 (2004) 788–793
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short Communication

Inaccurate reconstruction of ancestral GC levels creates a "vanishing isochores" effect☆

Fernando Alvarez-Valin,[a] Oliver Clay,[b] Stéphane Cruveiller,[b,c] and Giorgio Bernardi[b,*]

[a] *Secciòn Biomatemàtica, Facultad de Ciencias, Montevideo 11400, Uruguay*
[b] *Laboratory of Molecular Evolution, Stazione Zoologica "Anton Dohrn," Villa Comunale, 80121 Napoli, Italy*
[c] *Atelier de Génomique Comparative, Génoscope, Centre National de Séquençage, 2 rue Gaston Cremieux, CP 5706, 91057 Evry Cedex, France*

Available online at www.sciencedirect.com
ScienceDirect
Gene 385 (2006) 50–56
GENE
SECTION EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Are GC-rich isochores vanishing in mammals?

Jianying Gu [a,b], Wen-Hsiung Li [b,*]

[a] *Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, United States*
[b] *Department of Biology, College of Staten Island, City University of New York, Staten Island, NY 10314, United States*

Received 30 November 2005; accepted 30 March 2006
Available online 5 August 2006

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 31 (2004) 788–793
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short Communication

Inaccurate reconstruction of ancestral GC levels creates a "vanishing isochores" effect☆

Fernando Alvarez-Valin,[a] Oliver Clay,[b] Stéphane Cruveiller,[b,c] and Giorgio Bernardi[b,*]

[a] Secciòn Biomatemàtica, Facultad de Ciencias, Montevideo 11400, Uruguay
[b] Laboratory of Molecular Evolution, Stazione Zoologica "Anton Dohrn," Villa Comunale, 80121 Napoli, Italy
[c] Atelier de Génomique Comparative, Génoscope, Centre National de Séquençage, 2 rue Gaston Cremieux, CP 5706, 91057 Evry Cedex, France

Table 1
First-order analysis of 25 very GC_3-rich genes in cetartiodactyls[a]

Outgroup	Sequence name		Distance 1	Distance 2	Species average GC_3	A	B	C	Expected ratio GC → AT/ AT → GC
	Ingroup 1	Ingroup 2							
AF064555.PE1	BTPPI.PE1	OANIGFII1.PE1	0.115	0.336	92.7	0	2	2	7.608
AF181964	BT39469	Y13958	0.114	0.487	90.5	0	5	11	7.187
SSD158	BTCYB561.PE1	OAD157	0.032	0.148	90.2	0	1	3	2.577
AB038652.PMCT7	BTTRYPTMR.PE1	OAR18224.PE1	0.173	0.584	88.6	3	4	2	6.152
SSPROSDSN.PE1	BTAB4647.PE1	OAR133642.PE1	0.136	0.424	88.4	0	3	4	4.774
SSOXTRA.PE1	AF101724.PE1	OSOXYTREC.PE1	0.050	0.398	87.6	2	8	4	4.310
AF120326.PE1	AF074854	S44612.PE1	0.063	0.174	87.0	4	4	4	2.184
SSBLACMR.PE1	BTLGB.PE1	OALGB.PE1	0.087	0.316	85.0	3	3	2	2.853
SSCNP.CNP	BTCNP1.PE1	AF037467.CNP	0.055	0.207	84.7	0	3	0	2.146
SS12574.MYOD	AF093675	OAMYOD1.PE1	0.073	0.473	84.2	0	3	1	3.555
AF159382	S82652.PE1	AF034842	0.061	0.322	84.2	0	2	0	2.760
SSU59924.NOS	BTNOS.PE1	AF223471	0.108	0.326	84.1	6	4	7	2.698
U68482.G-CSF	AF092533.GCSF	OOCSFGR	0.106	0.225	83.0	0	5	2	2.000
SS14406.PE1	AF177290	OAPPCHY.PE1	0.073	0.241	82.9	2	5	4	2.119
SSGLUTP.PE1	BTGLUTI.GLUT-I	OAU89029.GLUT-1	0.100	0.247	82.3	7	11	8	2.056
U66254.OB	BT43943	OAU84247	0.073	0.215	81.1	0	3	2	1.834
AF064077	BTACTHA.PE1	AF116874.PE1	0.119	0.362	78.6	1	2	1	2.127
SSA005521	BTBRRIBO.PE1	S81745	0.047	0.384	78.4	0	2	0	2.225
SSTNFAB.PE2	AF011926.TNFA	OATNFA.PE1	0.037	0.339	78.4	1	1	1	2.099
SSJ001201.PE1	BTEP3B	AF035417	0.065	0.236	77.5	0	4	0	1.689
SS53020.PE1	BTY17260.STAR	S80098	0.089	0.224	77.1	2	3	0	1.621
SSIFNA1.PE1	BTIFNAA.PE1	OVU77908.PE1	0.078	0.332	76.7	4	1	0	1.906
SSMOTSA.PE1	AF068196.PE1	AF022771.PE1	0.042	0.249	76.5	0	2	0	1.692
SSINTL10A.IL-10	BT799.PE1	OA11421.IL-10	0.090	0.270	75.4	3	0	0	1.658
S96211.PE1	BTTIM.PE1	S67450.TIMP-1	0.069	0.236	75.2	0	5	0	1.577
						38	86	58	
			0.082	**0.310**	**82.8**				**3.22**

[a] Genes having a minimum $GC_3 > 75\%$ and represented by sequences in at least three cetartiodactyl species are shown in order of decreasing GC_3. The distance between the ingroup species (Distance 1) and the average distance between the ingroup species and the outgroup (Distance 2) were estimated by maximum likelihood using PAML (Yang, 2002). The table reports the GC_3 level of each gene in each species (indicated by its EMBL/GenBank name with ACNUC extension; retrieved from http://pbil.univ-lyon1.fr) as well as the mean GC_3 level across the three species. The expected ratio of (GC → AT)/(AT → GC) substitutions, at compositional equilibrium, was determined by the formulae of Eyre-Walker (1998). In calculating this expected value, we have, however, followed tradition in assuming constant substitution rates along a gene, and at CpG and non-CpG sites, although this first-order approximation tends to underestimate the expected value (see text). Genes and in-/outgroup species were taken from the supplementary table of Duret et al. (2002). A|B|C, number of synonymous substitutions from A or T to G or C | from G or C to A or T at non-CpG

Available online at www.sciencedirect.com
ScienceDirect
Gene 385 (2006) 50–56
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Are GC-rich isochores vanishing in mammals?

Jianying Gu [a,b], Wen-Hsiung Li [b,*]

[a] Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, United States
[b] Department of Biology, College of Staten Island, City University of New York, Staten Island, NY 10314, United States

R

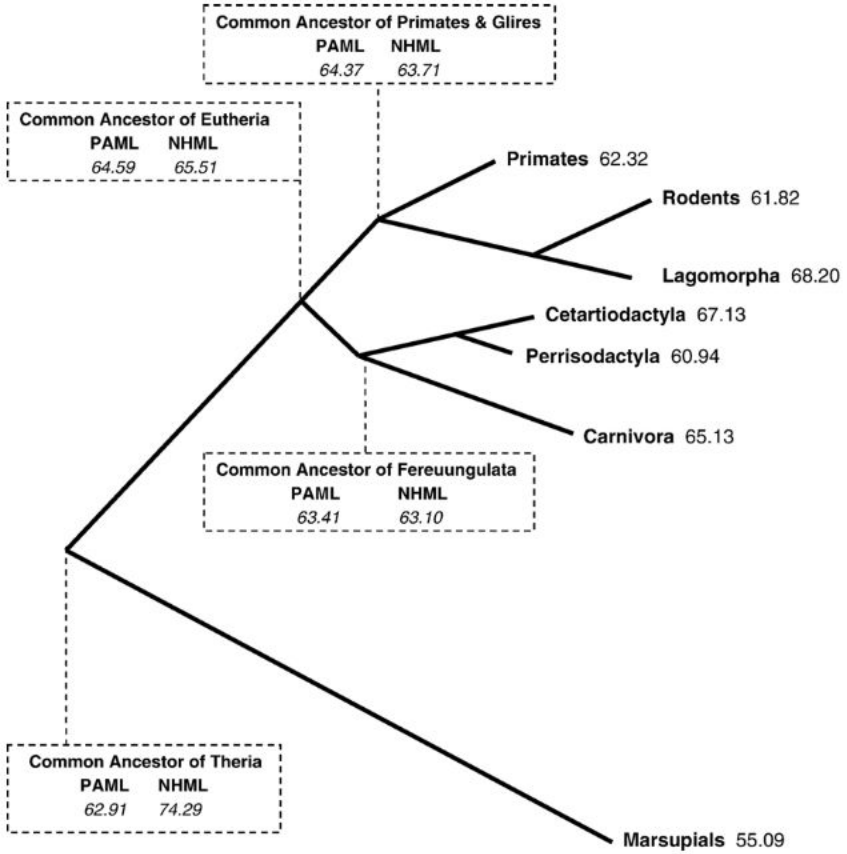


Fig. 1. Schematic phylogeny of mammals and inferred GC contents at codon position 3 (weighted average GC_3) at all internal and extant nodes. Two maximum likelihood methods (PAML and NHML) were applied to infer the ancestral state of nucleotides or GC content for the 176 genes with sequences from major eutherian groups and Metatheria. The GC_3 panels in the internal nodes represent the inferred GC content of the intermediate common ancestors. The GC_3 contents in each order studied and their common ancestral nodes were calculated using the sequence length as a weight. The branches of the schematic phylogeny are scaled to the average maximum likelihood estimates of each branch length of 176 genes.

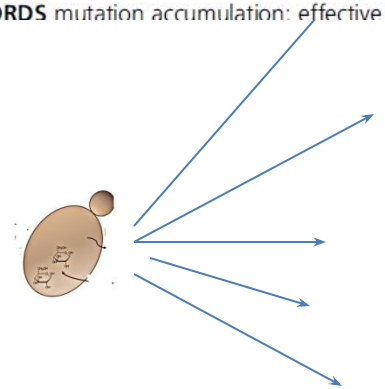
The Spontaneous Mutation Rate in the Fission Yeast *Schizosaccharomyces pombe*

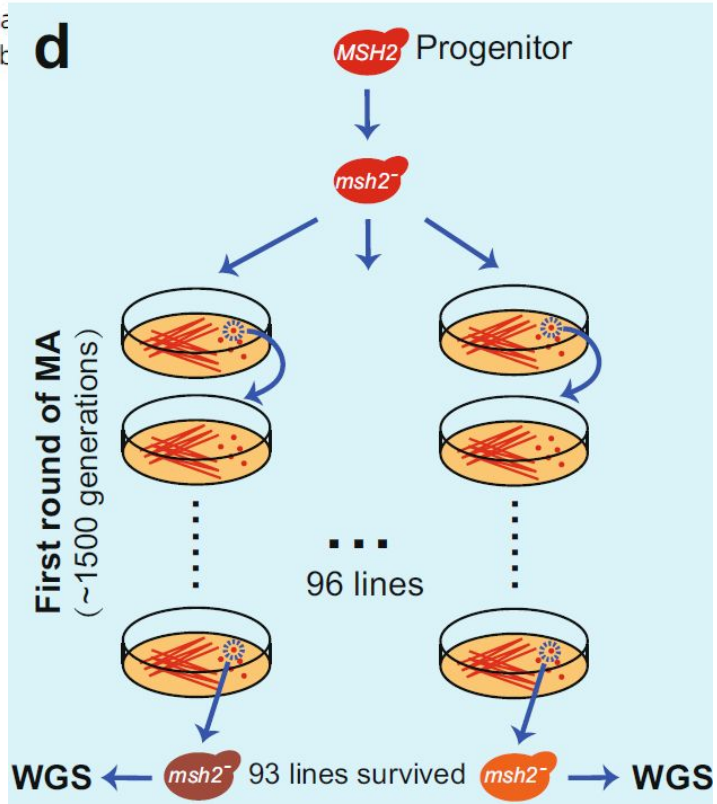
Ashley Farlow,*,†,1,2 Hongan Long,‡,1 Stéphanie Arnoux,* Way Sung,‡ Thomas G. Doak,‡,§ Magnus Nordborg,* and Michael Lynch‡,2

*Gregor Mendel Institute of Molecular Plant Biology, 1030 Vienna, Austria, †Institute of Population Genetics, 1210 Vetmeduni Vienna, Austria, and ‡Department of Biology and §School of Informatics and Computing, Indiana University, Bloomington, Indiana 47405

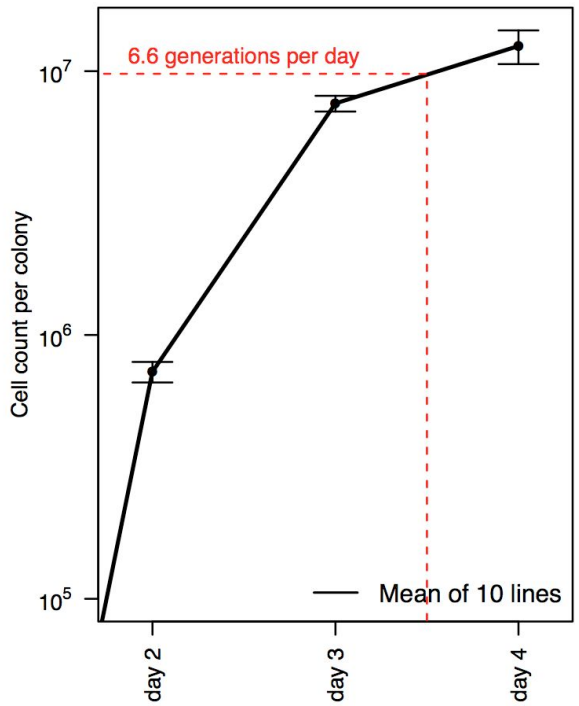
ABSTRACT The rate at which new mutations arise in the genome is a key factor in the evolution and adaptation of species. Here we describe the rate and spectrum of spontaneous mutations for the fission yeast *Schizosaccharomyces pombe*, a key model organism with many similarities to higher eukaryotes. We undertook an ~1700-generation mutation accumulation (MA) experiment with a haploid *S. pombe*, generating 422 single-base substitutions and 119 insertion-deletion mutations (indels) across the 96 replicates. This equates to a base-substitution mutation rate of 2.00×10^{-10} mutations per site per generation, similar to that reported for the distantly related budding yeast *Saccharomyces cerevisiae*. However, these two yeast species differ dramatically in their spectrum of base substitutions, the types of indels (*S. pombe* is more prone to insertions), and the pattern of selection required to counteract a strong AT-biased mutation rate. Overall, our results indicate that GC-biased gene conversion does not play a major role in shaping the nucleotide composition of the *S. pombe* genome and suggest that the mechanisms of DNA ma significantly between fission and budding yeasts. Unexpectedly, CpG sites appear to be excessively liab despite the likely absence of DNA methylation.

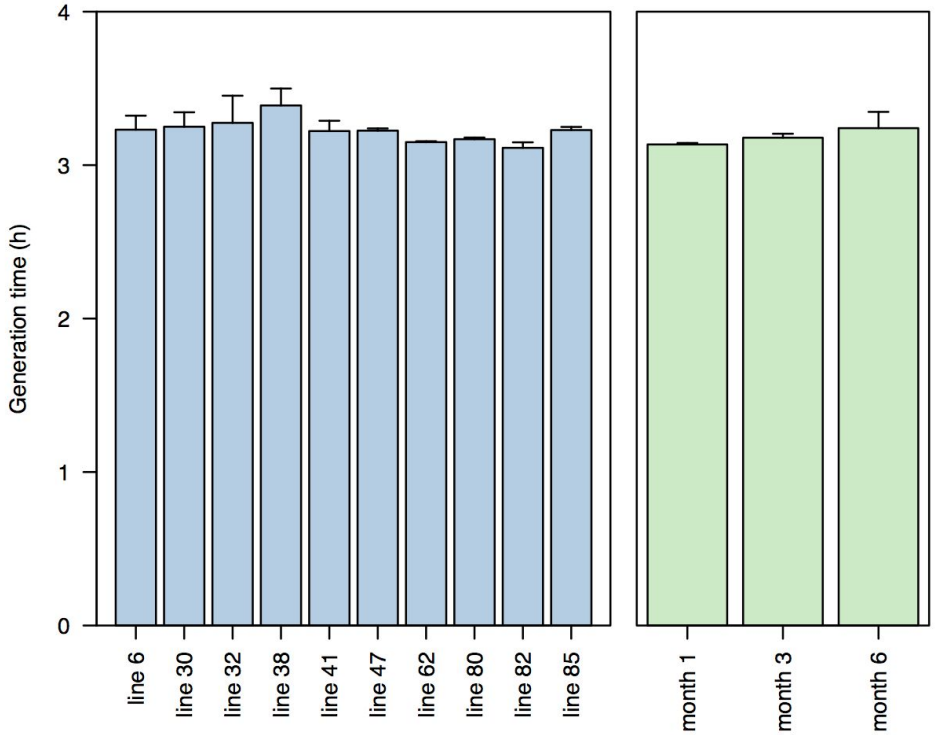
KEYWORDS mutation accumulation; effective population size; biased gene conversion; fission yeast





Numero de generaciones por hora

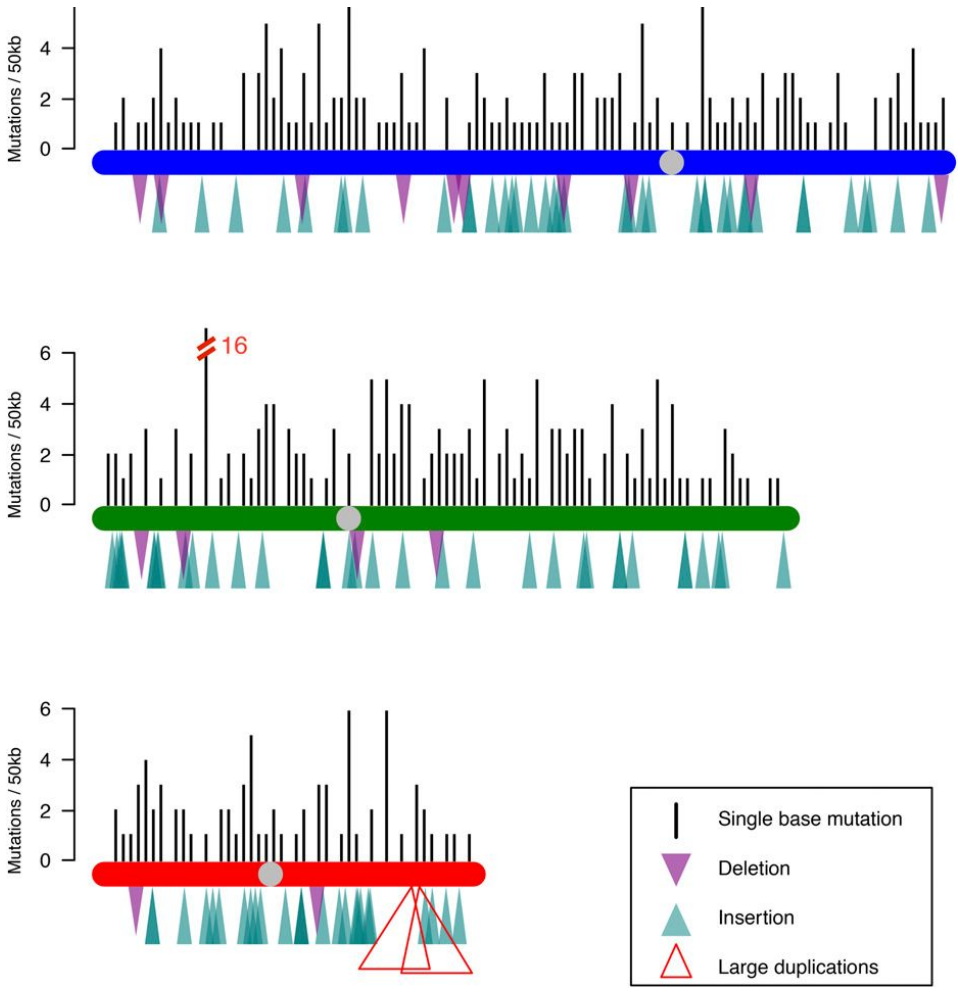




The Spontaneous Mutation Rate in the Fission Yeast *Schizosaccharomyces pombe*

Ashley Farlow,*,†,1,2 Hongan Long,‡,1 Stéphanie Arnoux,* Way Sung,‡ Thomas G. Doak,‡,§ Magnus Nordborg,* and Michael Lynch‡,2

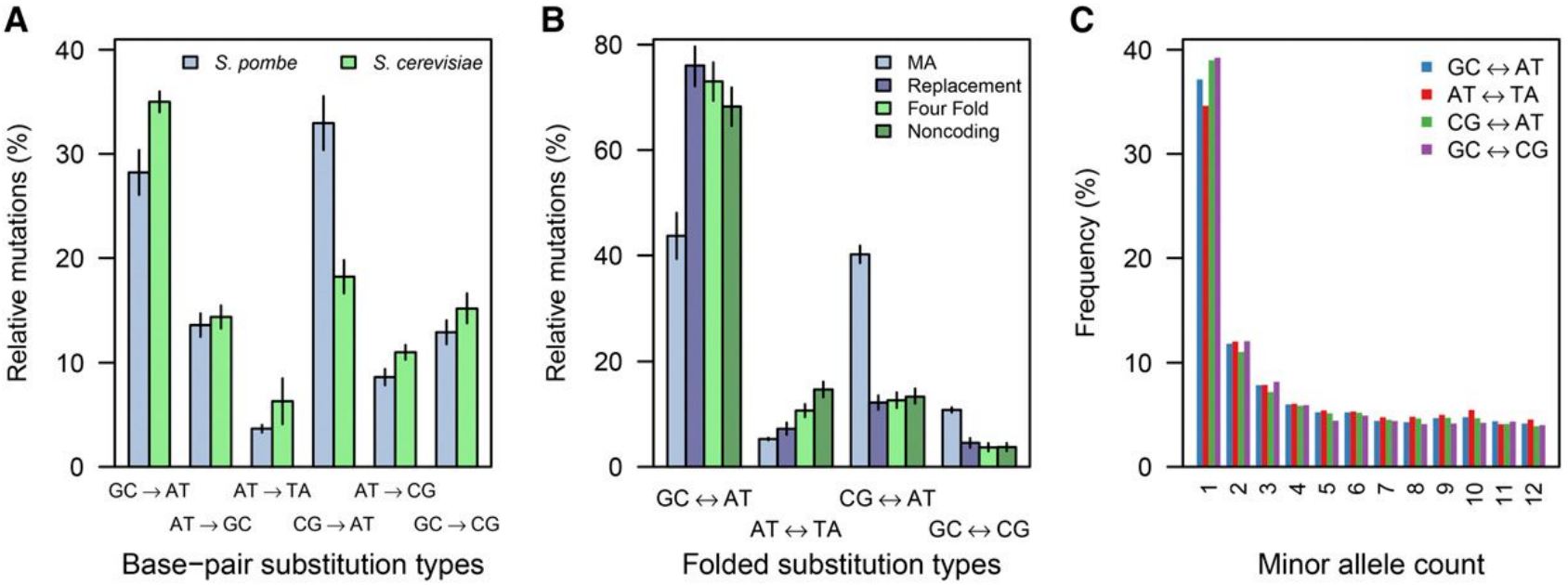
*Gregor Mendel Institute of Molecular Plant Biology, 1030 Vienna, Austria, †Institute of Population Genetics, 1210 Vetmeduni Vienna, Austria, and ‡Department of Biology and §School of Informatics and Computing, Indiana University, Bloomington, Indiana 47405



The Spontaneous Mutation Rate in the Fission Yeast *Schizosaccharomyces pombe*

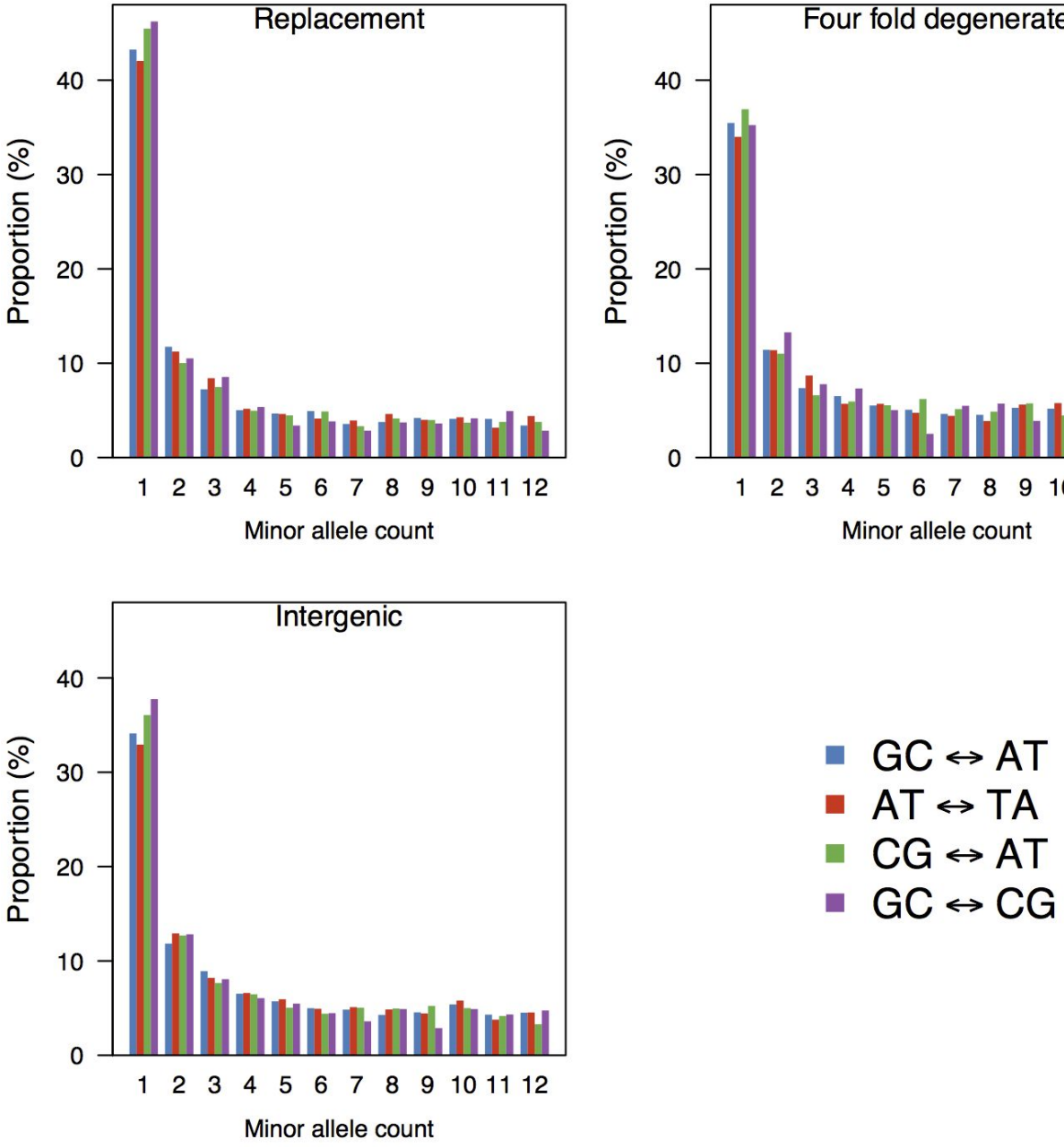
Ashley Farlow,*[,†,1,2] Hongan Long,[‡,1] Stéphanie Arnoux,* Way Sung,[‡] Thomas G. Doak,[‡,§] Magnus Nordborg,* and Michael Lynch[†,2]

*Gregor Mendel Institute of Molecular Plant Biology, 1030 Vienna, Austria, [†]Institute of Population Genetics, 1210 Vetmeduni Vienna, Austria, and [‡]Department of Biology and [§]School of Informatics and Computing, Indiana University, Bloomington, Indiana 47405



De novo

Comparacion 25 cepas


Replacement
Four fold degenerate
Intergenic
Proportion (%)
Minor allele count
0
10
20
30
40
1 2 3 4 5 6 7 8 9 10 11 12
GC ↔ AT
AT ↔ TA
CG ↔ AT
GC ↔ CG

The Spontaneous Mutation Rate in the Fission Yeast *Schizosaccharomyces pombe*

Ashley Farlow,*,†,1,2 Hongan Long,‡,1 Stéphanie Arnoux,* Way Sung,‡ Thomas G. Doak,‡,§ Magnus Nordborg,* and Michael Lynch‡,2

*Gregor Mendel Institute of Molecular Plant Biology, 1030 Vienna, Austria, †Institute of Population Genetics, 1210 Vetmeduni Vienna, Austria, and ‡Department of Biology and §School of Informatics and Computing, Indiana University, Bloomington, Indiana 47405

Seleccion para el contenido GC

The expected equilibrium A+T proportion of the genome $\tilde{p}$ is given by

$$\tilde{p} = \frac{v}{u+v}$$

where u is the A|T → G|C mutation rate (where | denotes "or"), and v is the G|C → A|T mutation rate (Lynch 2010). The population-scaled selection coefficient S favoring G|C over A|T is given by

$$p = \frac{me^S}{me^S + 1}$$

where p is the observed A+T proportion of the genome, and $m = v/u$.

De novo Mutations in Domestic Cat are Consistent with an Effect of Reproductive Longevity on Both the Rate and Spectrum of Mutations

Richard J. Wang [1,*] Muthuswamy Raveendran [2,3] R. Alan Harris [2,3] William J. Murphy [4] Leslie A. Lyons [5] Jeffrey Rogers [2,3] and Matthew W. Hahn [1,6]

[1]Department of Biology, Indiana University, Bloomington, IN, USA
[2]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA
[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, US
[4]Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA
[5]Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of
[6]Department of Computer Science, Indiana University, Bloomington, IN, USA

***Corresponding author:** E-mail: rjwang@indiana.edu.

Associate editor: Kelley Harris

Abstract

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-24364-6 OPEN

The rate and molecular spectrum of mutation are selectively maintained in yeast

Haoxuan Liu[1] & Jianzhi Zhang[1]✉

What determines the rate (μ) and molecular spectrum of mutation is a fundamental question. The prevailing hypothesis asserts that natural selection against deleterious mutations has pushed μ to the minimum achievable in the presence of genetic drift, or the drift barrier. Here ypothesis, μ substantially exceeds the drift barrier in diverse ccumulation (MA) in yeast frequently reduces μ, and deleting gene *PSP2* nearly halves μ. These results, along with a com- atural yeast strains, demonstrate that μ is maintained above election. Similar comparisons show that the mutation spec- T mutational bias is not intrinsic but has been selectively the separation of mutation from selection as distinct evolu- r to alleviating mutagenesis in various organisms by genome

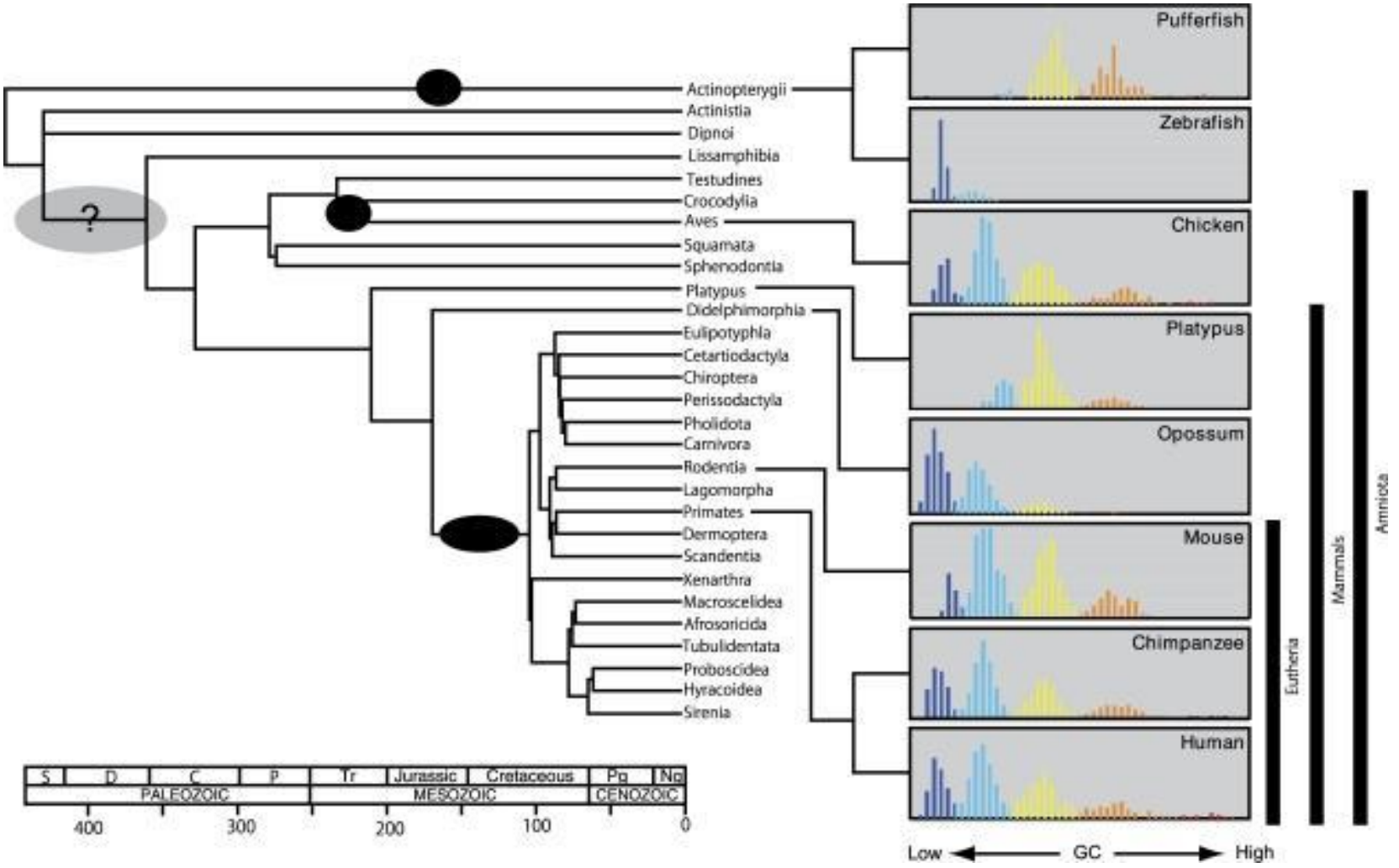
Article

Mutation bias reflects natural selection in *Arabidopsis thaliana*

https://doi.org/10.1038/s41586-021-04269-6
Received: 9 November 2020
Accepted: 17 November 2021
Published online: 12 January 2022
Open access
Check for updates

J. Grey Monroe[1,2]✉, Thanvi Srikant[1], Pablo Carbonell-Bejerano[1], Claude Becker[1,10], Mariele Lensink[2], Moises Exposito-Alonso[3,4], Marie Klein[1,2], Julia Hildebrandt[1], Manuela Neumann[1], Daniel Kliebenstein[2], Mao-Lun Weng[5], Eric Imbert[6], Jon Ågren[7], Matthew T. Rutter[8], Charles B. Fenster[9] & Detlef Weigel[1]✉

Since the first half of the twentieth century, evolutionary theory has been dominated by the idea that mutations occur randomly with respect to their consequences[1]. Here we test this assumption with large surveys of de novo mutations in the plant *Arabidopsis thaliana*. In contrast to expectations, we find that mutations occur less often in functionally constrained regions of the genome—mutation frequency is reduced by half inside gene bodies and by two-thirds in essential genes. With independent genomic mutation datasets, including from the largest *Arabidopsis* mutation accumulation experiment conducted to date, we demonstrate that


Actinopterygii
Actinistia
Dipnoi
Lissamphibia
Testudines
Crocodylia
Aves
Squamata
Sphenodontia
Platypus
Didelphimorphia
Eulipotyphla
Cetartiodactyla
Chiroptera
Perissodactyla
Pholidota
Carnivora
Rodentia
Lagomorpha
Primates
Dermoptera
Scandentia
Xenarthra
Macroscelidea
Afrosoricida
Tubulidentata
Proboscidea
Hyracoidea
Sirenia
?
Pufferfish
Zebrafish
Chicken
Platypus
Opossum
Mouse
Chimpanzee
Human
Eutheria
Mammals
Amniota
S
D
C
P
Tr
Jurassic
Cretaceous
Pg
Ng
PALEOZOIC
MESOZOIC
CENOZOIC
400
300
200
100
0
Low
GC
High