Selective Constraints on Amino Acids Estimated by a Mechanistic Codon Substitution Model with Multiple Nucleotide Changes

Sanzo Miyazawa

miyazawa@smlab.sci.gunma-u.ac.jp

Graduate School of Engineering, Gunma University, Japan

presented at

The 48th annual meeting of Biophysical Society of Japan in 2010 in Sendai.

(September 20-22, 2010)

ABSTRACT

Background: Here, we develop a new codon-based model, which consists of mutations at the nucleotide level and selection at the amino acid level via a genetic code, and estimate selective constraints on amino acids from available empirical substitution matrices. The mutational process of individual codons is modeled as a reversible Markov process, and multiple nucleotide changes are assumed to occur with the same order of time as single nucleotide changes do. In a codon substitution model, 1829 codon exchangeabilities must be determined. In the present model, they are expressed as functions of 6 nucleotide mutation rates and selective constraints for 190 types of amino acid replacements. If the selective constraints are estimated, and their relative strengths among amino acid replacements are approximated to be constant irrespective of protein families, parameters to be optimized in maximum likelihood (ML) and Bayesian inferences of phylogenetic trees will be drastically reduced to parameters for nucleotide mutations and some additional ones.

Results: The present model with substitution rates that are assumed to obey a Γ distribution can be well fitted to each 1-PAM matrix of empirical amino acid substitution matrices (JTT, WAG, and LG) and empirical codon substitution matrix (KHG) already published. ML estimators of selective constraints on amino acids are calculated together with other parameters. Akaike information criterion (AIC) values indicate that the assumption of multiple nucleotide changes significantly better fits the model to the empirical substitution matrices. One of interesting results is that the ML estimators of transition to transversion bias obtained from these empirical matrices are not so large as previously estimated. Also, the present model with the selective constraints estimated from the JTT/WAG/LG/KHG can be well fitted to other matrices including the ones (cpREV) for chloroplast proteins and (mtREV) for vertebrate mitochondrial proteins.

Conclusions: Thus, the present codon-based model with the ML estimators for the selective constraints and with adjustable mutation rates of nulceotides would be useful as a simple substitution model in ML and Bayesian inferences of molecular phylogenetic trees, and enables us to to obtain biologically meaningful information at both nucleotide and amino acid levels from codon and protein sequences.

1. INTRODUCTION

Purpose and distinctive features of the present study:

- To develop a mechanistic codon-based substitution model for protein evolution.
  - A time-homogeneous and reversible Markov model, which consists of mutations at the nucleotide level and selection at the amino acid level.
  - Codon substitutions due to multiple nucleotide changes are taken into account.
- To confirm the significance of multiple nucleotide changes in codon substitutions.
- To estimate selective constraints for all types of amino acid pairs.
- To confirm that the present codon model with the estimate of selective constraints is a better substitution model than empirical substitution matrices such as the JTT; it allows to estimate mutational parameters at the nucleotide level from homologous coding sequences and to obtain better likelihoods in a probabilistic inference of their phylogenetic relations.

2. METHODS

A time-homogeneous and reversible Markov model for codon substitutions

Transition matrix over time t: $S(t) = \exp(Rt)$ with $f_\mu R_{\mu\nu} = f_\nu R_{\nu\mu}$

Substitution rate matrix: $R_{\mu\nu} = \text{const}\, M_{\mu\nu} \frac{f_\nu}{f_\nu^{\text{mut}}} e^{w_{\mu\nu}}$ for $\mu \neq \nu$, normalized to $\sum_\mu f_\mu R_{\mu\mu} = -1$

Mutation rate matrix: $M_{\mu\nu} = \prod_{i=1}^{3} [\delta_{\mu_i\nu_i} + (1 - \delta_{\mu_i\nu_i}) m_{\mu_i\nu_i} f_{i,\nu_i}^{\text{mut}}]$ for $\mu \neq \nu$

Selective constraints: $e^{w_{\mu\nu}} = \sum_a \sum_b C_{\mu a} C_{\nu b} e^{w_{ab}}$

where $\mu = (\mu_1, \mu_2, \mu_3), \nu = (\nu_1, \nu_2, \nu_3) \in \{$ codons $\}$, $\mu_i, \nu_j \in \{t, c, a, g\}$,and $a, b \in \{$amino acids$\}$.

f_μ	Equilibrium frequency of codon μ
f_μ^{mut}	Equilibrium frequency of codon μ for the M; $f_{\mu=(\mu_1,\mu_2,\mu_3)}^{\text{mut}} = f_{\mu_1}^{\text{mut}} f_{\mu_2}^{\text{mut}} f_{\mu_3}^{\text{mut}}$
$f_{\mu_i}^{\text{mut}}$	Equilibrium frequency of nucleotide μ_i for the M
$w_{ab} = w_{ba}$	Selective constraint against substitutions between amino acids a and b; $w_{aa} = 0$ and $w_{ab} < 0$ for $a \neq b$
$C_{\mu a}$	Genetic code table; $C_{\mu a} = 1$ if μ is a codon for amino acid a, otherwise 0
$m_{\mu_i\nu_i}$	Exchangeability between nucleotides μ_i and ν_i; $m_{\mu_i\nu_i} = m_{\nu_i\mu_i}$

Likelihood of an empirical codon/amino acid substitution matrix

Log-likelihood: $\ell(\boldsymbol{\theta}) = N \sum_\kappa \sum_\lambda f_\kappa^{\text{obs}} S_{\kappa\lambda}^{\text{obs}} \log(f_\kappa \langle S \rangle(\tau,\sigma)_{\kappa\lambda})$ $\kappa, \lambda = \mu, \nu$ or a, b.

Kullback-Leibler Information: $\hat{I}_{\text{KL}}(\boldsymbol{\theta}) = -\ell(\boldsymbol{\theta}) + \sum_\kappa \sum_\lambda f_\kappa^{\text{obs}} S_{\kappa\lambda}^{\text{obs}} \log(f_\kappa^{\text{obs}} S_{\kappa\lambda}^{\text{obs}})$ $\kappa, \lambda = \mu, \nu$ or a, b.

Mean of $S(t)$ over t or rate: $\langle S \rangle(\tau,\sigma) = \int_0^\infty S(t)\Gamma(t;\tau,\sigma)dt = [(I - \sigma R)^{-1}]^\tau$

$\mu = (\mu_1, \mu_2, \mu_3)$, $\mu_i \in \{t, c, a, g\}$, $a, b \in \{\text{amino acids}\}$

Estimation of parameters:

Amino acid frequencies: $\hat{f}_a = f_a^{\text{obs}}$

Codon frequencies: $\hat{f}_\mu = f_\mu^{\text{obs}}$ or $C_{\mu a}\hat{f}_\mu = f_a^{\text{obs}} C_{\mu a} \hat{f}_\mu^{\text{usage}} / \sum_\nu C_{\nu a} \hat{f}_\nu^{\text{usage}}$, $\hat{f}^{\text{usage}}_{\mu=(\mu_1,\mu_2,\mu_3)} = \hat{f}^{\text{usage}}_{\mu_1} \hat{f}^{\text{usage}}_{\mu_2} \hat{f}^{\text{usage}}_{\mu_3}$

Shape parameter $\hat{\tau}$ of Γ: $\sum_\kappa \hat{f}_\kappa \langle S(\hat{\tau},\sigma) \rangle_{\kappa\kappa} = \sum_\kappa f_\kappa^{\text{obs}} S_{\kappa\kappa}^{\text{obs}}$ $\kappa, \lambda = \mu, \nu$ or a, b.

By maximizing a likelihood: $\hat{\boldsymbol{\theta}} \equiv \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}} \hat{I}_{\text{KL}}(\boldsymbol{\theta})$

- Relative exchangeabilities: $m_{tc|ag}/m_{[tc][ag]}$, $m_{ag}/m_{tc|ag}$, $m_{ta}/m_{[tc][ag]}$, $m_{tg}/m_{[tc][ag]}$, $m_{ca}/m_{[tc][ag]}$
- Relative ratio of multiple nucleotide changes: $m_{[tc][ag]}$
- Scale parameter of Γ : σ of a Γ distribution for rate variation
- Selective constraints: $\hat{w}_{ab}$ estimated by maximizing a likelihood (ML) or
  $\hat{w}_{ab} = \hat{\beta} w_{ab}^{\text{estimate}} + \hat{w}_0(1 - \delta_{ab})$; $\hat{\beta}$ and $\hat{w}_0$ are estimated by ML.
- Codon usage: $f_{\mu_i}^{\text{usage}}$ if codon frequencies are unkown.

Evaluation of model: **Smaller AIC or Δ AIC means a better model.**

Akaike Information Criterion: $\mathrm{AIC} \equiv -2\ell(\hat{\boldsymbol{\theta}}) + 2 \cdot (\text{number of adjustable parameters})$

$$\Delta\mathrm{AIC} \equiv \mathrm{AIC} + 2N \sum_\kappa \sum_\lambda f_\kappa^{\mathrm{obs}} S_{\kappa\lambda}^{\mathrm{obs}} \log(f_\kappa^{\mathrm{obs}} S_{\kappa\lambda}^{\mathrm{obs}}) = 2N \hat{I}_{\mathrm{KL}}(\boldsymbol{\theta}) + 2 \cdot (\#\text{parameters})$$

Log-odds: $\text{log-}O(\langle S \rangle(t))_{\kappa\lambda} \equiv \dfrac{10}{\log 10} \log \dfrac{\langle S \rangle(t)_{\kappa\lambda}}{f_\lambda}$ $\kappa, \lambda = \mu, \nu$ or a, b.

Empirical substitution matrices used for model fitting:

The following 1 PAM matrices are used.

JTT amino acid:	compiled from closely related proteins by Jones et al. (1992).
cpREV amino acid:	estimated from chloroplast proteins (Adachi et al., 2000),
mtREV amino acid:	estimated from vertebrate mitochondrial proteins (Adachi & Hasegawa, 1996) by maximizing the likelihood of a given phylogenetic tree.
WAG amino acid:	estimated from proteins encoded in nuclear DNA (Whelan &Goldman, 2001),
LG amino acid:	estimated from proteins encoded in nuclear DNA (Le &Gascuel, 2008),
KHG codon:	estimated from protein-coding sequences in nuclear DNA (Kosiol et al., 2007) by maximizing the likelihood of given phylogenetic trees and branch lengths.

3. RESULTS

Table 1.

Model name	Brief description
ML-n	Selective constraints $\{w_{ab}\}$ are estimated by maximizing the likelihood of an empirical substitution matrix. The suffix n means the number of ML parameters.
ML-87	In the ML-87, multiple nucleotide changes are disallowed, and $\{w_{ab}\}$ for all 75 single-step amino acid pairs are estimated.
ML-91	In the ML-91, multiple nucleotide changes are allowed, and $\{w_{ab}\}$ for all 75 single-step amino acid pairs and for 6 groups of multiple-step amino acid pairs are estimated. Equal codon usage is assumed.
ML-200	In the ML-200 for codon substitution matrices, $\{w_{ab}\}$ for all 190 amino acid pairs are estimated.
ML-n+	First, the ML-n is used to estimate parameters, and then $\{w_{ab}\}$ for all multiple-step amino acid pairs are estimated by maximizing the likelihood with fixing all other parameters at the values estimated by the ML-n.
JTT-ML91-n, WAG-ML91-n, LG-ML91-n	$w_{ab} = \beta w_{ab}^{\text{JTT/WAG/LG-ML91}} + w_0(1-\delta_{ab})$, where $w_{ab}^{\text{JTT/WAG/LG-ML91}}$ is one estimated by maximizing the likelihodd of the JTT/WAG/LG$_{\text{amino acid}}$ in the ML-91. The suffix n means the number of ML parameters.
JTT-ML91+-n, WAG-ML91+-n, LG-ML91+-n	$w_{ab} = \beta w_{ab}^{\text{JTT/WAG/LG-ML91+}} + w_0(1-\delta_{ab})$, where $w_{ab}^{\text{JTT/WAG/LG-ML91+}}$ is one estimated from the JTT/WAG/LG$_{\text{amino acid}}$ in the ML-91+. The suffix n means the number of ML parameters. The JTT/WAG/LG-ML91+-0 correspond to the JTT/WAG/LG-F.
KHG-ML200-n	$w_{ab} = \beta w_{ab}^{\text{KHG-ML200}} + w_0(1-\delta_{ab})$, where $w_{ab}^{\text{KHG-ML200}}$ is one estimated by maximizing the likelihodd of the KHG$_{\text{codon}}$ in the ML-200. The suffix n means the number of ML parameters. The KHG-ML200-0 correspond to the KHG-F.

The effects of multiple nucleotide changes in the 1-PAM JTT

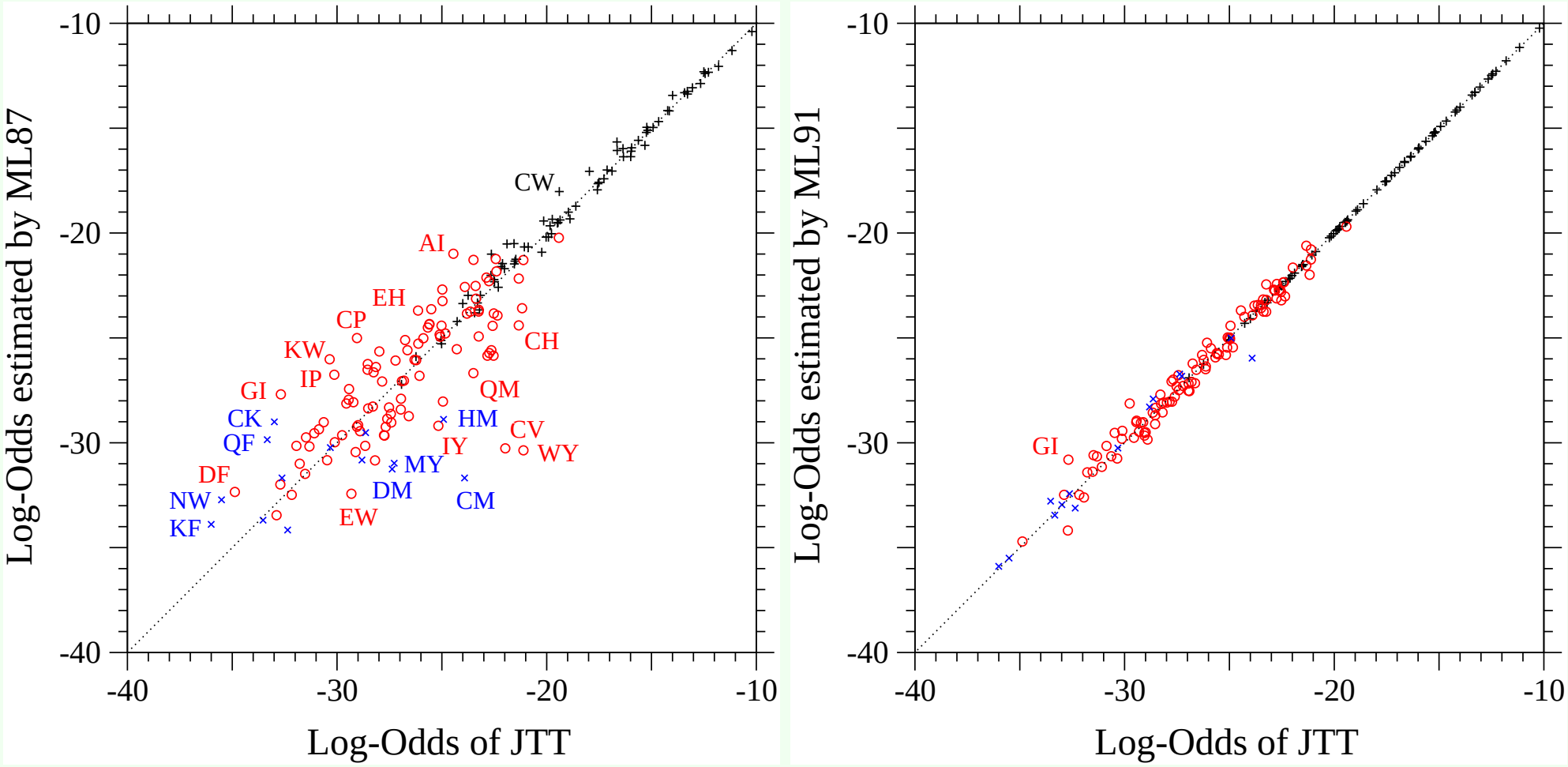
- AIC is significantly improved by taking account of multiple nucleotide changes.

(A) ML-87: Single nucleotide changes only

ΔAIC = 2072.0

(B) ML-91: Multiple nucleotide changes allowed

ΔAIC = 297.5



+ single nucleotide change o double nucleotide change x triple nucleotide change

Fig. 1

Table 2.

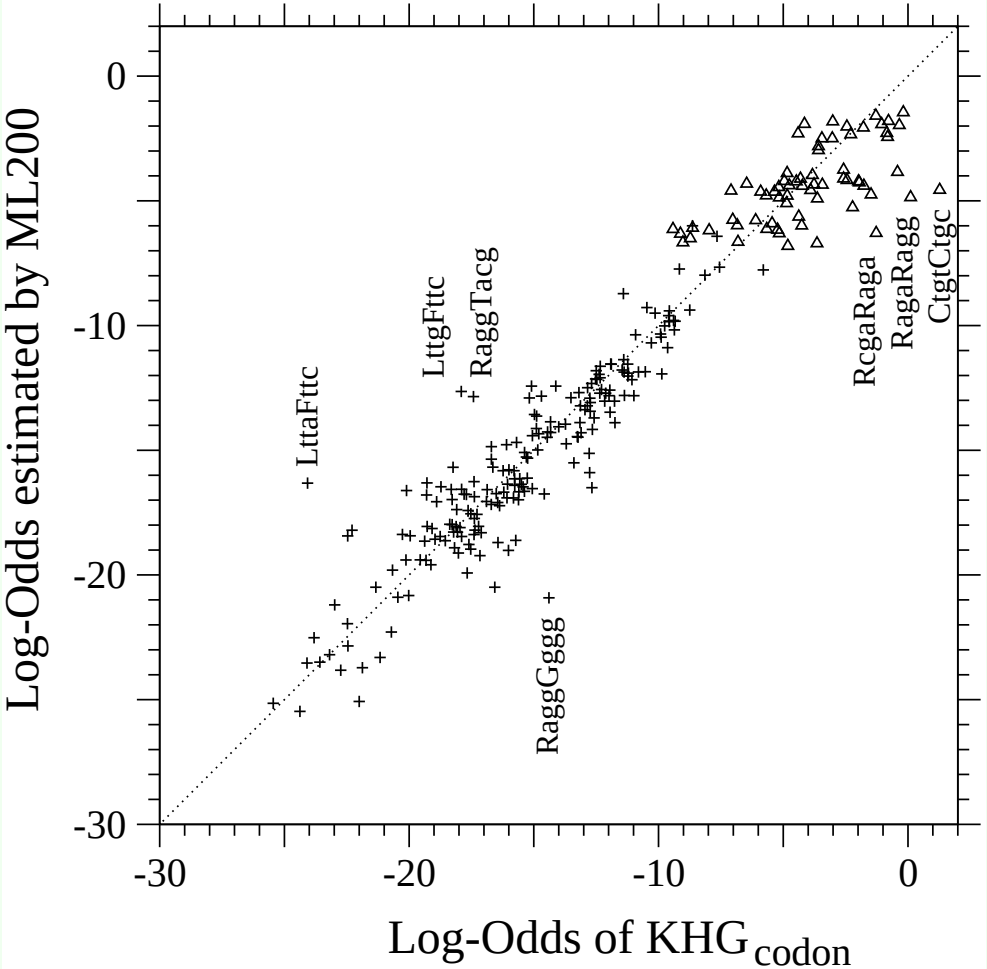
ML estimators in the present models fitted to empirical substitution matrices

id		JTT		WAG		LG	KHG (codon)
no.	parameter	ML–87	ML–91	ML–87	ML–91	ML–91	ML–200
0	$-\hat{w}_0$	N/A	N/A	N/A	N/A	N/A	N/A
1	$1/\hat{\beta}$	N/A	N/A	N/A	N/A	N/A	N/A
2	$\hat{m}_{[tc][ag]}$	$(\to 0)$	0.637	$(\to 0)$	1.28	1.08	0.939
3	$\hat{m}_{tc\|ag}/\hat{m}_{[tc][ag]}$	0.0919	1.57	0.746	1.70	1.85	0.843
4	$\hat{m}_{ag}/\hat{m}_{tc\|ag}$	1.77	1.14	1.98	1.32	1.23	0.945
5	$\hat{m}_{ta}/\hat{m}_{[tc][ag]}$	0.0293	0.729	0.0477	0.791	0.676	1.52
6	$\hat{m}_{tg}/\hat{m}_{[tc][ag]}$	3.21	0.940	3.64	1.04	1.07	0.554
7	$\hat{m}_{ca}/\hat{m}_{[tc][ag]}$	0.719	1.19	0.110	1.23	1.28	0.573
8	$\hat{f}^{\mathrm{mut}}_{t+a}$	0.408	0.459	0.372	0.367	0.388	0.497
9	$\hat{f}^{\mathrm{mut}}_{t}/\hat{f}^{\mathrm{mut}}_{t+a}$	0.113	0.501	0.234	0.587	0.450	0.513
10	$\hat{f}^{\mathrm{mut}}_{c}/\hat{f}^{\mathrm{mut}}_{c+g}$	0.698	0.429	0.425	0.479	0.427	0.470
11	$\hat{f}^{\mathrm{usage}}_{t+a}$	0.0682	(0.5)	0.0669	(0.5)	(0.5)	NA
12	$\hat{f}^{\mathrm{usage}}_{t}/\hat{f}^{\mathrm{usage}}_{t+a}$	0.461	(0.5)	0.330	(0.5)	(0.5)	NA
13	$\hat{f}^{\mathrm{usage}}_{c}/\hat{f}^{\mathrm{usage}}_{c+g}$	0.386	(0.5)	0.310	(0.5)	(0.5)	NA
14	$\hat{\sigma}$	27.3	0.738	43.3	0.905	0.415	$\to 0$
	$\hat{\tau}\hat{\sigma}$	0.334	0.0243	0.317	0.0223	0.0246	0.0240
	#parameters	107	111	107	111	111	261
	$\hat{I}_{KL}(\hat{\boldsymbol{\theta}}) \times 10^8$ [a]	15695	638	35319	1903	2771	269946
	ΔAIC [b]	2072.0	297.5	1370.8	284.3	782.5	unknown
Ratio of substitution rates per codon							
	the total base/codon	1.28	1.35	1.38	1.53	1.38	1.29 (1.29)[c]
	transition/transversion	0.464	1.08	0.482	0.932	1.18	0.764 (0.765)[c]
	nonsynonymous/synonymous[d]	1.13	1.37	1.57	2.07	1.05	0.726 (0.723)[c]
Ratio of substitution rates per codon for $\sigma \to 0$							
	total base/codon	1.0	1.22	1.0	1.38	1.31	1.29
	transition/transversion	0.101	1.21	0.647	1.11	1.31	0.764
	nonsynonymous/synonymous[d]	0.0644	1.04	0.138	1.50	0.853	0.726
Ratio of substitution rates per codon for $w_{ab} = 0$ and $\sigma \to 0$							
	total base/codon	1.0	1.45	1.0	1.72	1.67	1.51
	transition/transversion	0.0605	0.829	0.499	0.933	0.992	0.427
	nonsynonymous/synonymous[d]	11.3	5.58	11.1	8.68	7.45	6.81

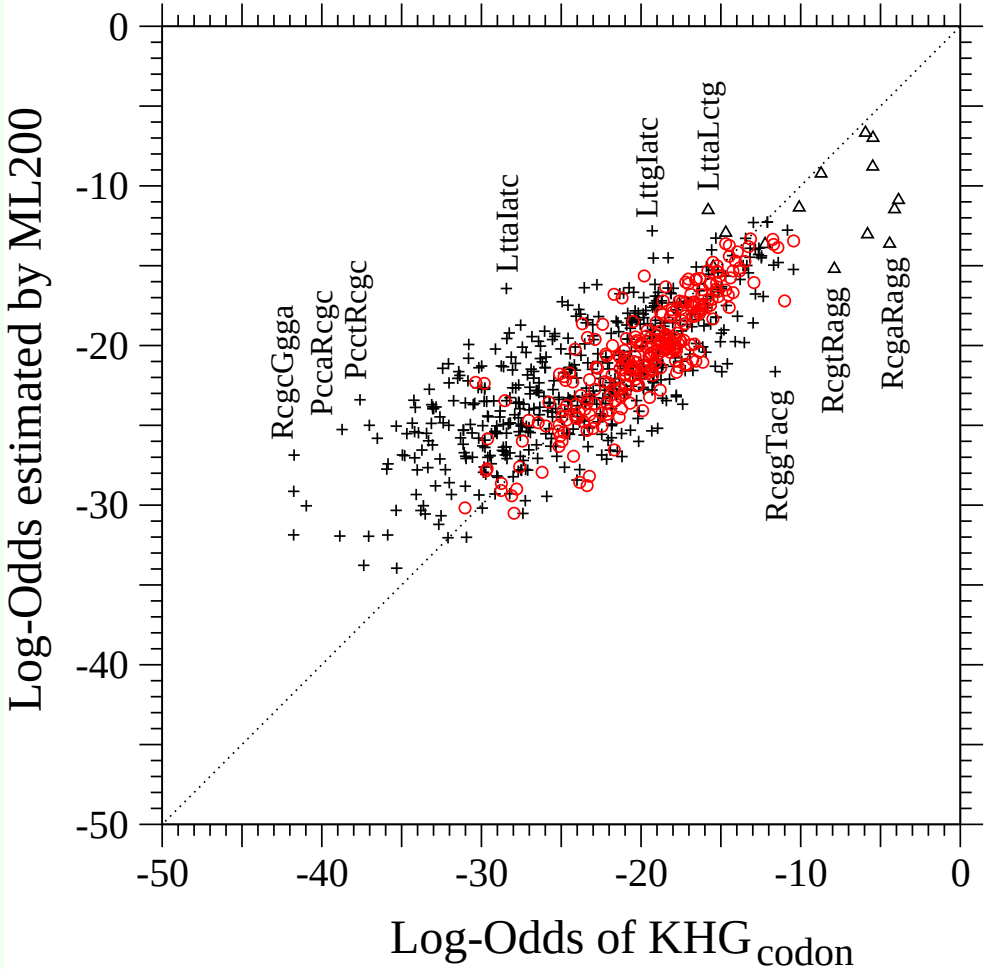
Comparison of the observed and the expected log-odds in the 1-PAM KHG_{codon}

ML-200 model: The selective constraints ($\hat{w}_{ab}$) for all 190 amino acid pairs are optimized.

(A) Codon pairs of single nucleotide changes

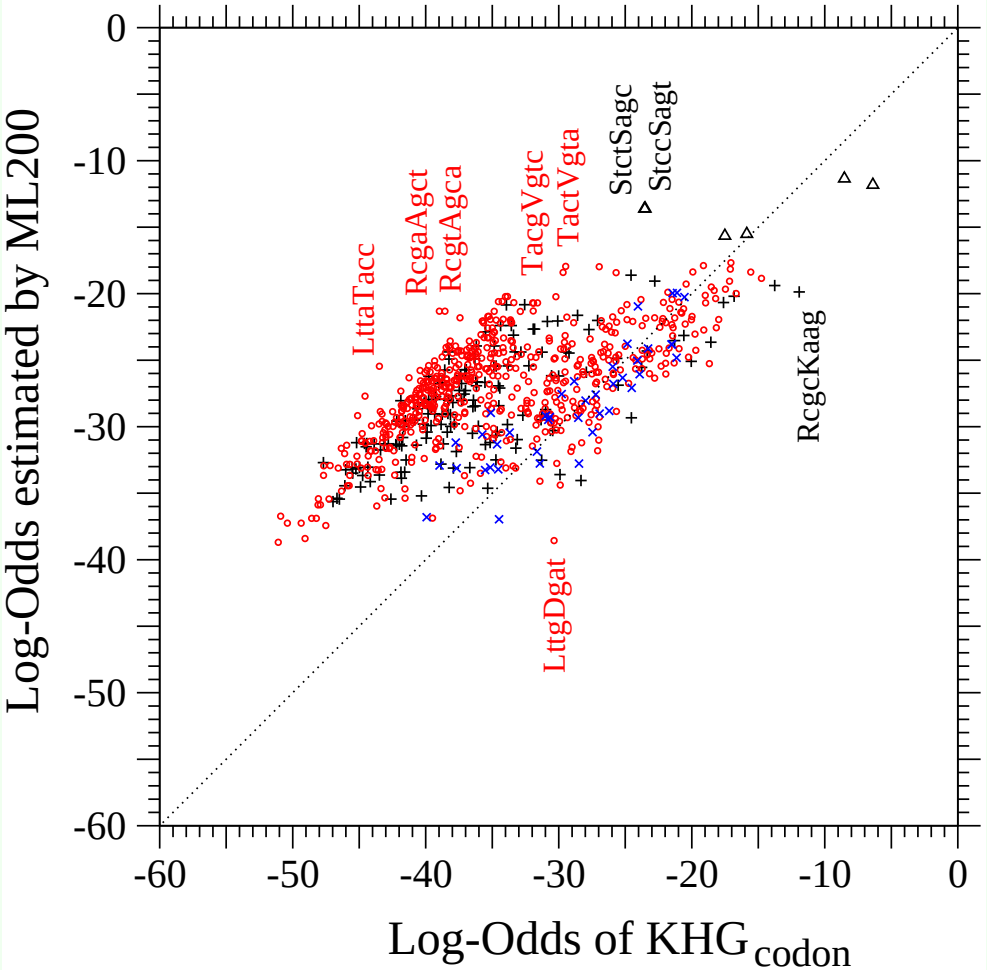


(B) Codon pairs of double nucleotide changes

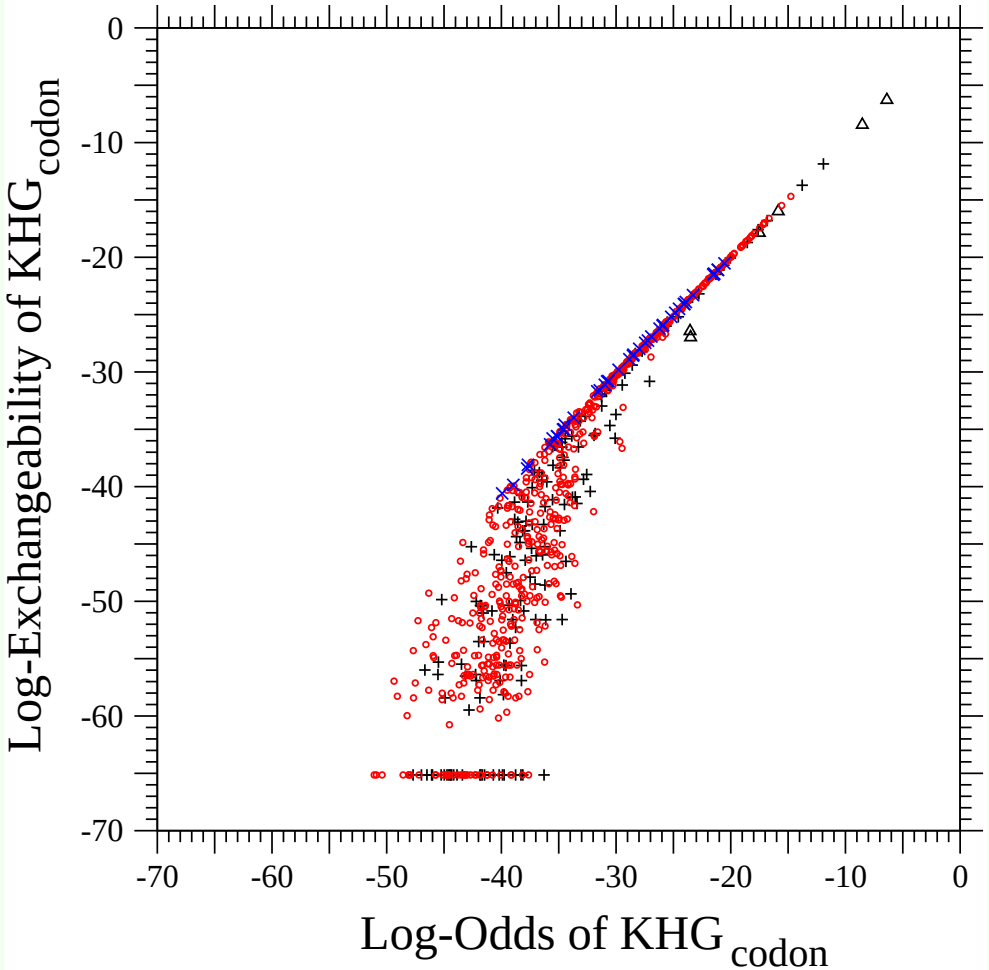


△ synonymous amino acid pair + one-step amino acid pair o two-step amino acid pair

(C) Codon pairs of triple nucleotide changes



(D) Log-exchangeabilities of triple nucleotide changes



△ synonymous amino acid pair + one-step amino acid pair o two-step amino acid pair x three-step amino acid pair

Fig. 2

Correlation of selective constraints ($\hat{w}_{ab}$) between various estimates

Table 3.

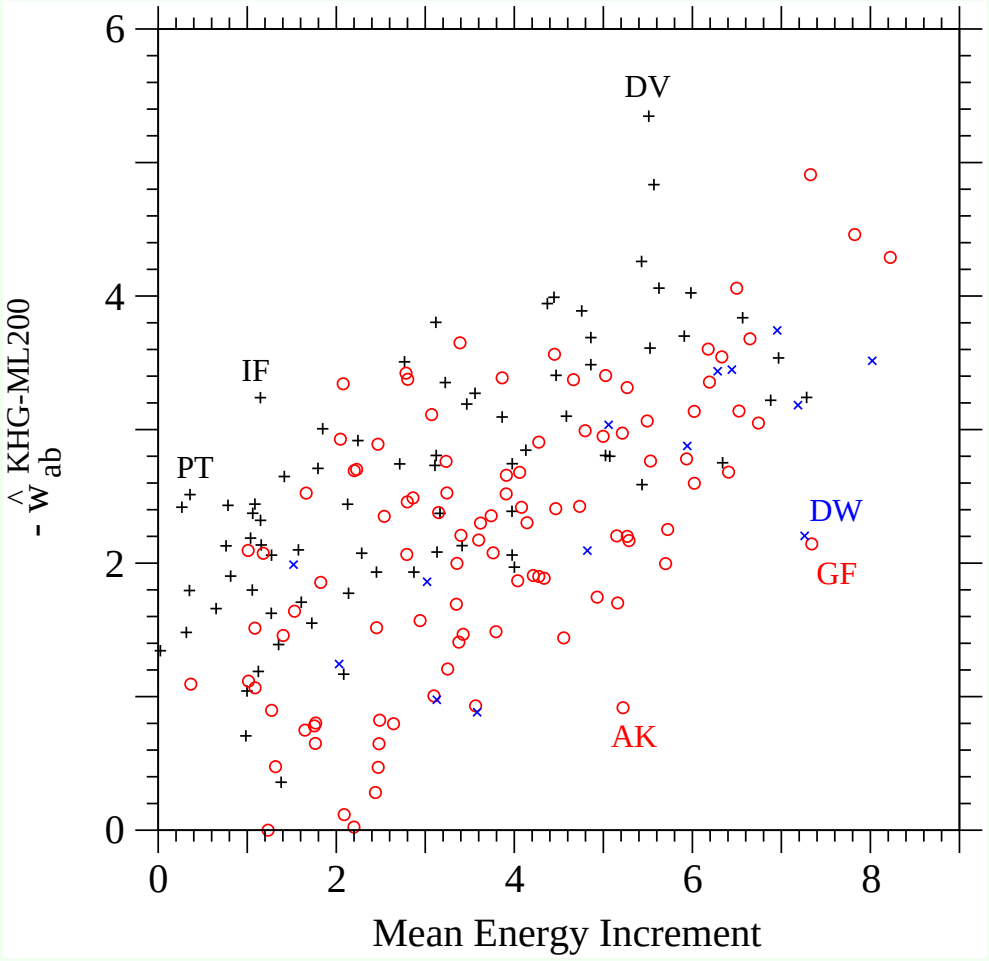
	EI	JTT-ML91+	WAG-ML91+	LG-ML91+	KHG-ML200
EI		0.45	0.51	0.59	0.55 (0.65)[a]
JTT-ML91+	0.66		0.80	0.80	0.51
WAG-ML91+	0.68	0.87		0.86	0.55
LG-ML91+	0.71	0.82	0.90		0.58
KHG-ML200	0.71	0.77	0.69	0.74	

Upper half: Correlation coefficients for 75 single-step amino acid pairs.

Lower half: for 86 multi-step amino acid pairs; 29 pairs of the least exchangeable category is excluded.

EI: Physico-chemical estimates of selective constraints based on residue-residue contact energies.

$\hat{w}_{ab}$ (KHG-ML200 vs. EI)



+ single nucleotide change o double nucleotide change x triple nucleotide change

Correlation coefficients = 0.60 (0.71, 0.63, 0.79)

Mean Energy Increment (EI): due to an amino acid replacement, based on residue-residue contact energies.

Fig. 3

Performance of various estimates ($\hat{w}_{ab}$) of selective constraints

- $w_{ab} = \beta w_{ab}^{\text{estimate}} + w_0(1 - \delta_{ab})$, where $w^{\text{estimate}} \equiv \hat{w}^{\text{JTT/WAG/LG-ML91+}}$ or $\hat{w}^{\text{KHG-ML200}}$.
- Parameters including β and w_0 are optimized.

Table 4.

Model No.	#parameters (id no. [a])	ΔAIC [b]					$\hat{I}_{KL}(\hat{\boldsymbol{\theta}}) \times 10^8$ [c]	
		JTT	WAG	LG	cpREV	mtREV	KHG (amino acid)	KHG (codon)
JTT-ML91+								
0	20		2657.5	20807.0	412.1	426.0		
11	31(1-10,14)		1152.9	12140.0	264.5	286.5	40931	
12	32(0-10,14)							473668
WAG-ML91+								
0	20	9095.4		10537.3	283.7	535.1		
11	31(1-10,14)	3299.2		4813.3	235.9	365.1	12789	
12	32(0-10,14)							496804
LG-ML91+								
0	20	13669.8	1806.0		434.5	593.4		
11	31(1-10,14)	3878.5	574.7		243.0	314.9	5732	
12	32(0-10,14)							436557
KHG-ML200								
0	20	15063.5	953.4	12568.9	360.8	593.6		
11	31(1-10,14)	4429.9	518.7	3006.1	229.4	334.1		

[a] Parameter id numbers in the parenthesis mean ML parameters in each model and other parameters are fixed to the value of the corresponding parameter listed in the column of the ML-91 or the ML-200 in Table 2; each id number corresponds to the parameter id number listed in Table 2.

[b] $\Delta\text{AIC} \equiv 2N\hat{I}_{KL}(\hat{\boldsymbol{\theta}}) + 2\times$ #parameters with $N \simeq 5919000$ for the JTT, $N \approx 1637663$ for the WAG, $N \approx 10114373$ for the LG, $N \approx 149355$ for the cpREV, and $N \approx 137637$ for the mtREV; see text for details.

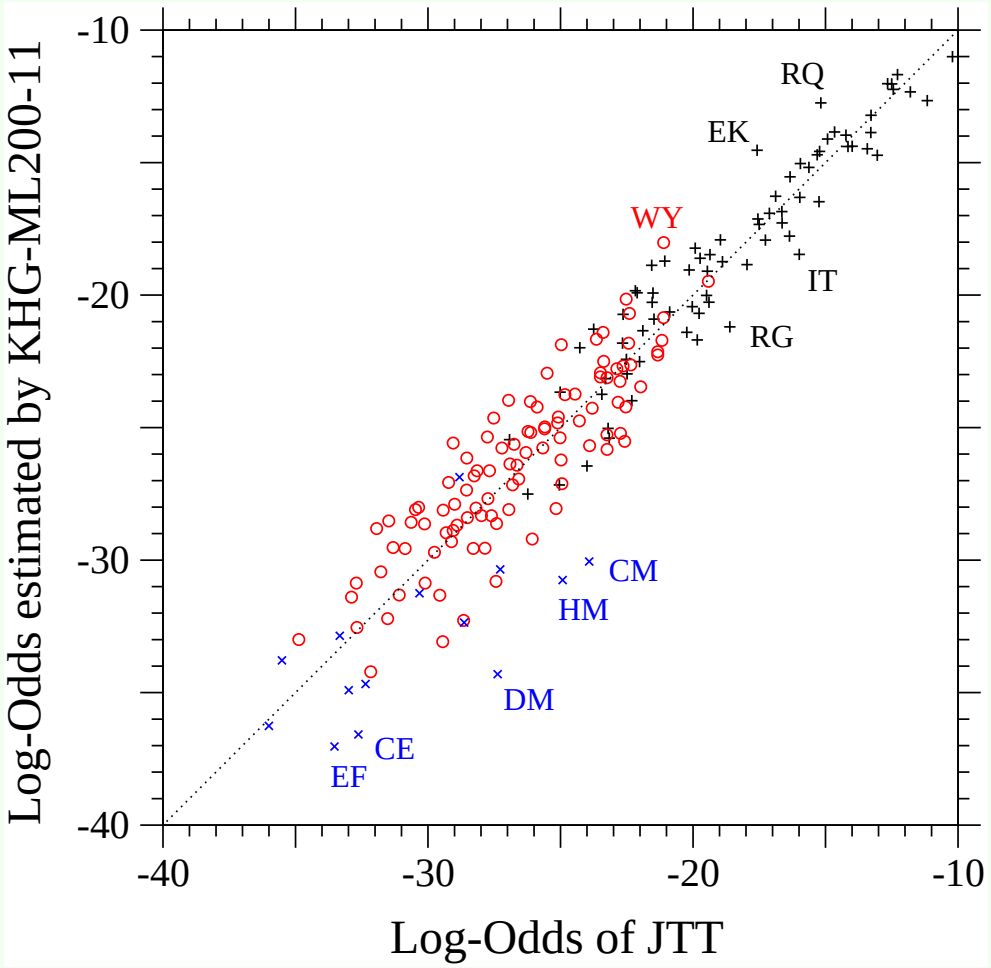
[c] $\hat{I}_{KL}(\hat{\boldsymbol{\theta}}) = -(\ell(\hat{\boldsymbol{\theta}})/N + 2.97009788)$ for the KHG-derived amino acid substitution probability matrix, and $-(\ell(\hat{\boldsymbol{\theta}})/N + 4.19073314)$ for the KHG codon substitution probability matrix; see text for details.

Performance of $\hat{w}^{\text{KHG-ML200}}$ estimated from the $\text{KHG}_{\text{codon}}$ in the ML-200

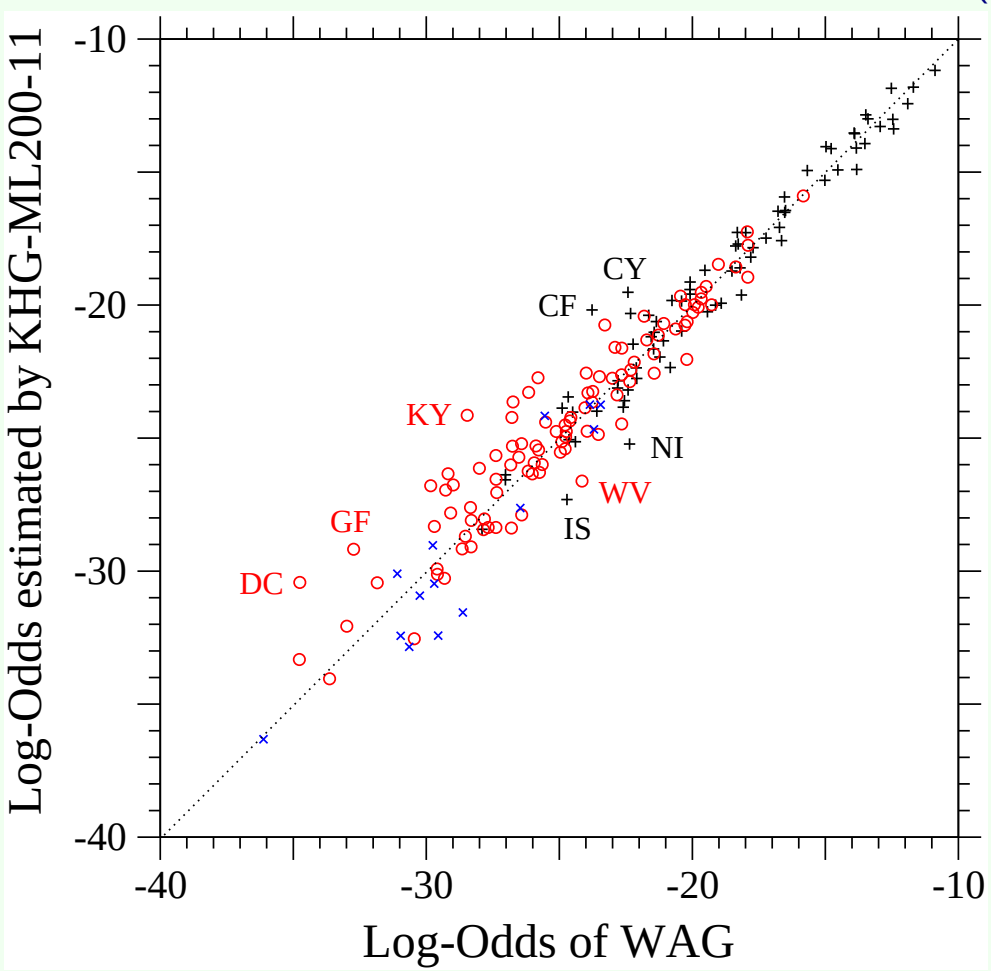
- $w_{ab} = \beta \hat{w}_{ab}^{\text{KHG-ML200}} + w_0(1 - \delta_{ab})$; all parameters including β and w_0 are optimized.

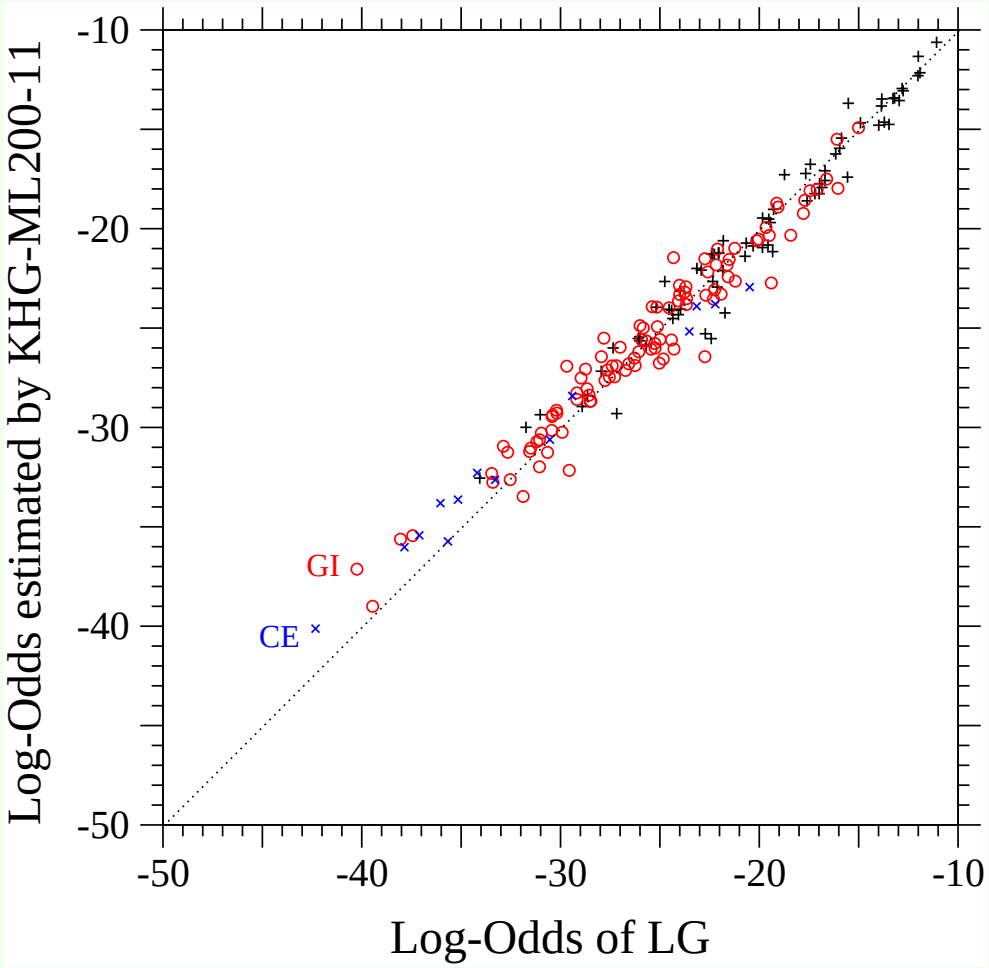
Empirical substitution matrices can be well reproduced.

(A)

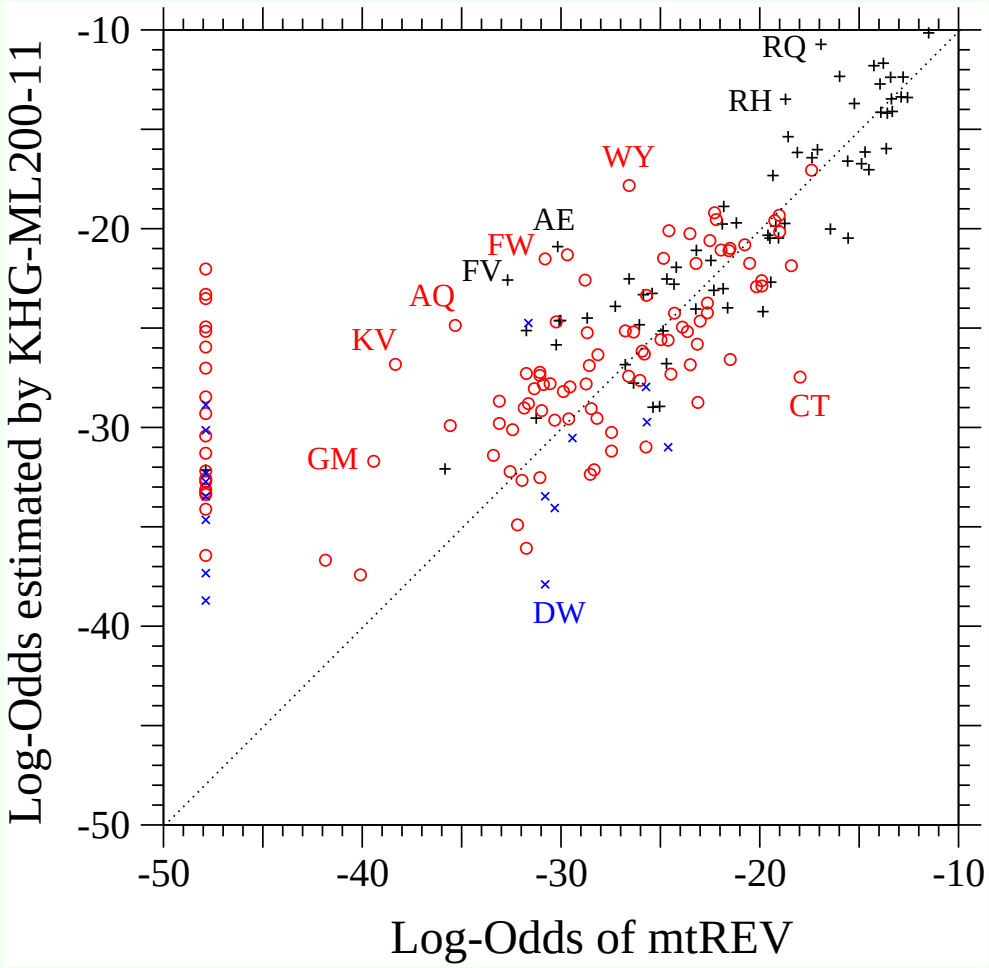


(B)




Log-Odds estimated by KHG-ML200-11
Log-Odds of LG
-10
-20
-30
-40
-50
GI
CE

(C)


Log-Odds estimated by KHG-ML200-11
Log-Odds of mtREV
-10
-20
-30
-40
-50
RQ
RH
WY
AE
FW
FV
AQ
KV
GM
CT
DW

(D)

Fig. 4

AICs of a phylogenetic tree of the concatenated coding sequences of 12 proteins in mtDNA from 20 vertebrate species.

- The AIC of the tree is significantly improved by the present codon model with $w^{\text{KHG-ML200}}$.

Codon Substitution Model	#c[a]	#parameters	ℓ	AIC	Description
mtREV-F[b]	1	59	-121053.2	242224.4	$\alpha = \infty^a$
	4	60	-120183.4	240486.7	$\hat{\alpha} = 0.704^a$
KHG-ML200-11-F	1	59	-118010.8	236139.7	$\alpha = \infty^a$, $(\sigma = 2.89, w_0 = 0^e)$
fitted to mtREV[c]	4	60	-117281.4	234682.8	$\hat{\alpha} = 3.86^a$, $(\sigma = 2.89, w_0 = 0^e)$
KHG-ML200-12-F[d]	1	71	-115176.0	230494.0	$\alpha = \infty^a$, $\hat{\sigma} = 0.496$, $\hat{w}_0 = -1.46$, $\hat{\beta} = 1.04$, $\hat{m}_{[tc][ag]} = 0.242$, $\hat{m}_{ta,tg,ca}/\hat{m}_{[tc][ag]} = 1.72, 0.591, 0.797$, $\hat{m}_{tc\|ag}/\hat{m}_{[tc][ag]} = 2.48$, $\hat{m}_{ag}/\hat{m}_{tc\|ag} = 0.830$, $\hat{f}^{\text{mut}}_{a,c,g} = 0.242, 0.272, 0.229$
	4	72	-113336.0	226816.1	$\hat{\alpha} = 1.15^a$, $\hat{\sigma} = 0.00$, $\hat{w}_0 = -1.53$, $\hat{\beta} = 1.20$, $\hat{m}_{[tc][ag]} = 0.155$, $\hat{m}_{ta,tg,ca}/\hat{m}_{[tc][ag]} = 1.65, 0.643, 0.645$, $\hat{m}_{tc\|ag}/\hat{m}_{[tc][ag]} = 2.61$, $\hat{m}_{ag}/\hat{m}_{tc\|ag} = 0.839$, $\hat{f}^{\text{mut}}_{a,c,g} = 0.263, 0.294, 0.200$

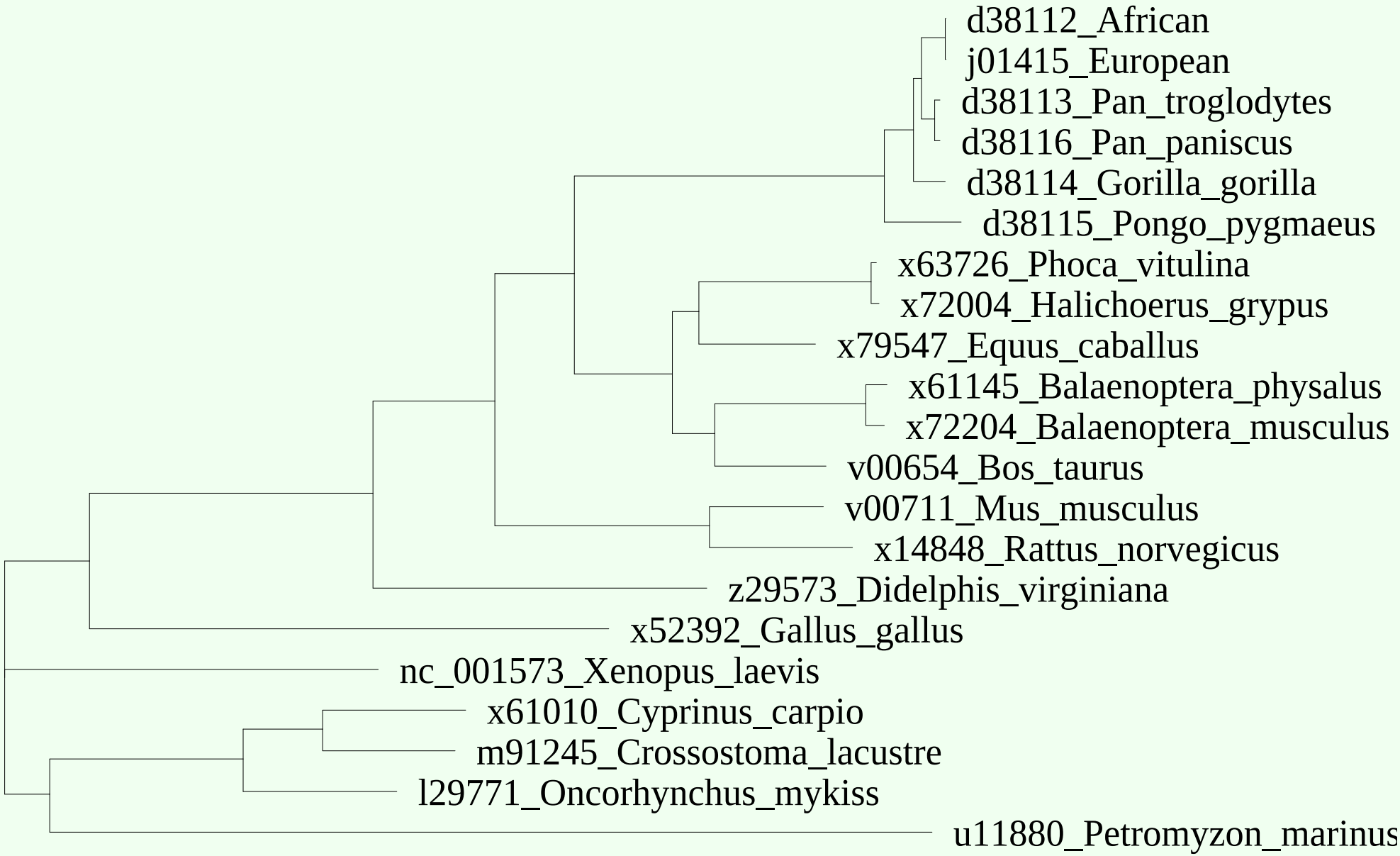
[a] The number of substitution rate categories for a discrete Γ distribution. α is the shape parameter.

[b] Exchangeabilties for nonsynonymous codons are equal to those of the corresponding amino acid pairs in the mtREV , and those for synonymous codons are equal to $-\infty$.

[c] The codon substitution matrix is that of the KHG-ML200-11 fitted to the mtREV.

[d] The ML estimators $\{w^{\text{KHG-ML200}}\}$ are used and all other parameters except codon frequencies are estimated by maximizing the likelihood of the tree.

Phylogenetic tree of mtDNA from 20 vertebrate species.



4. CONCLUSION

- Codon substitutions due to multiple nucleotide changes are significant, and must be taken into account to model the substitution process of amino acids.
- The present model with the ML estimate of selective constraints is a better substitution model with a few adjustable parameters at the nucleotide level than empirical substitution matrices.
  - it can well reproduce empirical substitution matrices such as the JTT/WAG/LG/KHG/mtREV/cpREV.
  - it can yield better AICs of phylogenetic trees than empirical substitution matrices.
  - it allows to estimate mutational tendencies at the nucleotide level in phylogenetic analyses of protein coding sequences; a transition to transversion rate ratio, a non-synonymous to synonymous rate ratio and the ratio of multiple nucleotide changes.