

Supplemental Material

Amino acid composition in endothermic vertebrates is biased in the same direction as in thermophilic prokaryotes

Guang-Zhong Wang & Martin J. Lercher

Supplemental Figures	2
Supplemental Tables.....	5

Supplemental Figures

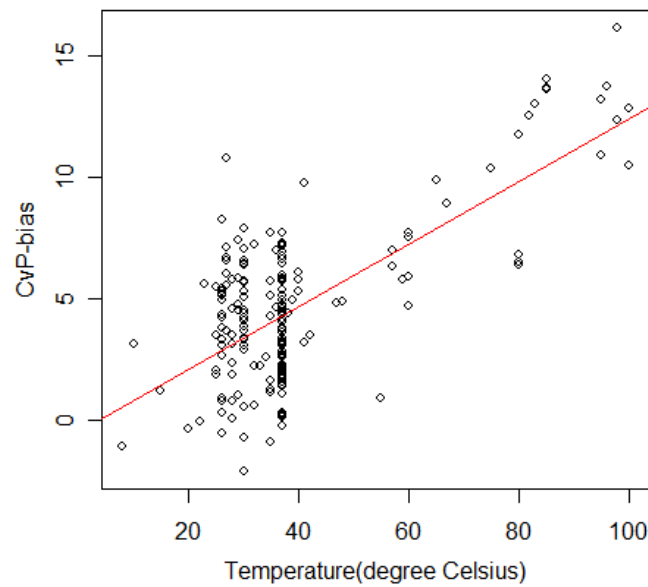


Figure S1. Correlation between an alternative measure of amino acid bias related to protein stability, $CvP\text{-}bias = D + E + R + K - N - Q - T - S$, and optimal growth temperature in 204 prokaryotes (Pearson's $R = 0.69$, $p < 10^{-15}$; Spearman's $\rho = 0.31$, $p = 5 \times 10^{-6}$).

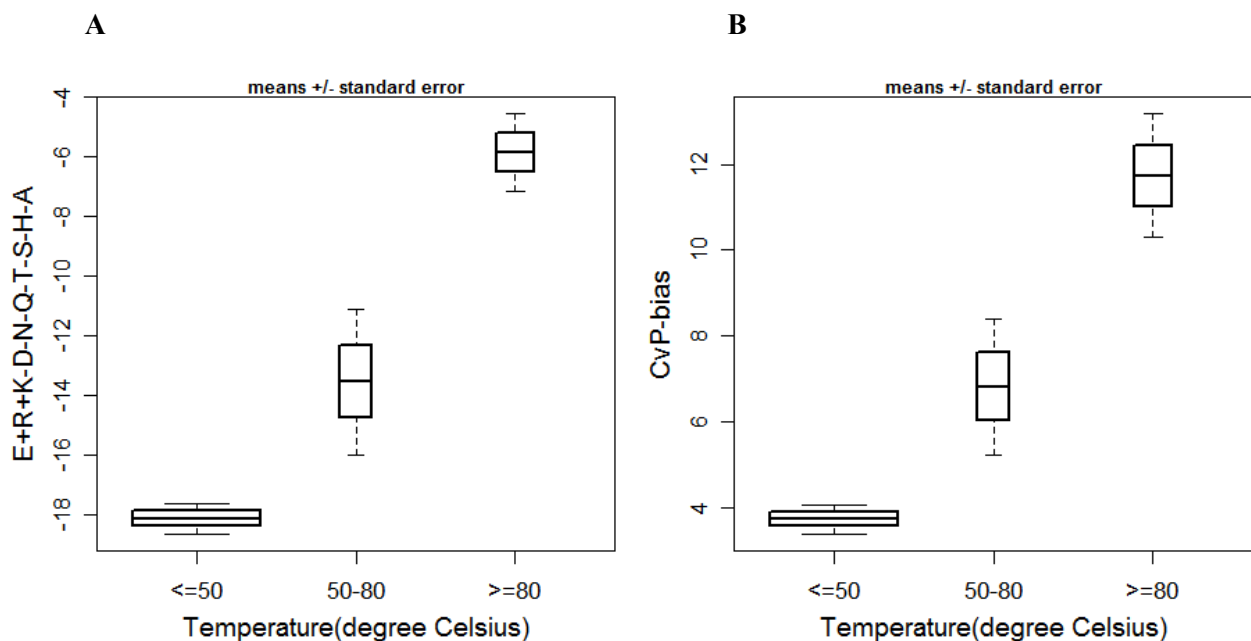


Figure S2. Differences between measures of amino acid bias related to protein stability and optimal growth temperature in hyperthermophiles (optimal growth temperature $OGT \geq 80^\circ\text{C}$), thermophiles ($OGT = 50\text{--}80^\circ\text{C}$), and mesophiles ($OGT \leq 50^\circ\text{C}$), using either $ERK = E + R + K - D - N - Q - T - S - H - A$ (**A**) or $CvP\text{-}bias = D + E + R + K - N - Q - T - S$ (**B**) as measures of amino acid bias. All pairwise comparisons show highly significant differences (Wilcoxon rank sum tests for ERK ($CvP\text{-}bias$): $p = 2 \times 10^{-5}$ (0.00023) between hyperthermophiles and thermophiles, $p = 0.00059$ (0.00024) between thermophiles and mesophiles, and $p = 4 \times 10^{-11}$ (2×10^{-10}) between hyperthermophiles and mesophiles). Central lines are means, boxes give standard error, whiskers 2x standard error.

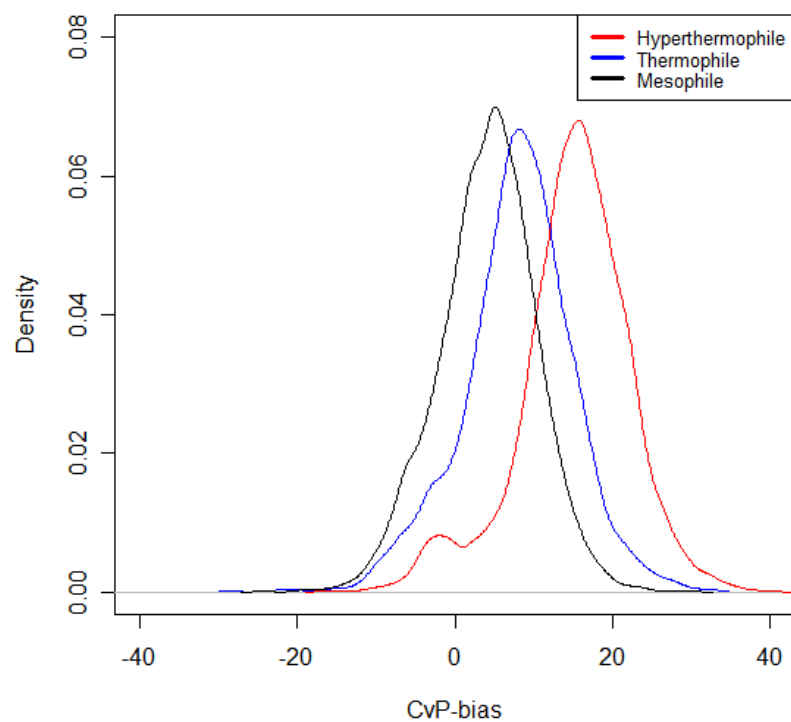


Figure S3. Distribution of *CvP-bias* across orthologous proteins for 5 species each of hyperthermophile (optimal growth temperature $\geq 80^{\circ}$ Celsius), thermophile ($50-80^{\circ}$), and mesophile ($\leq 50^{\circ}$) prokaryotes.

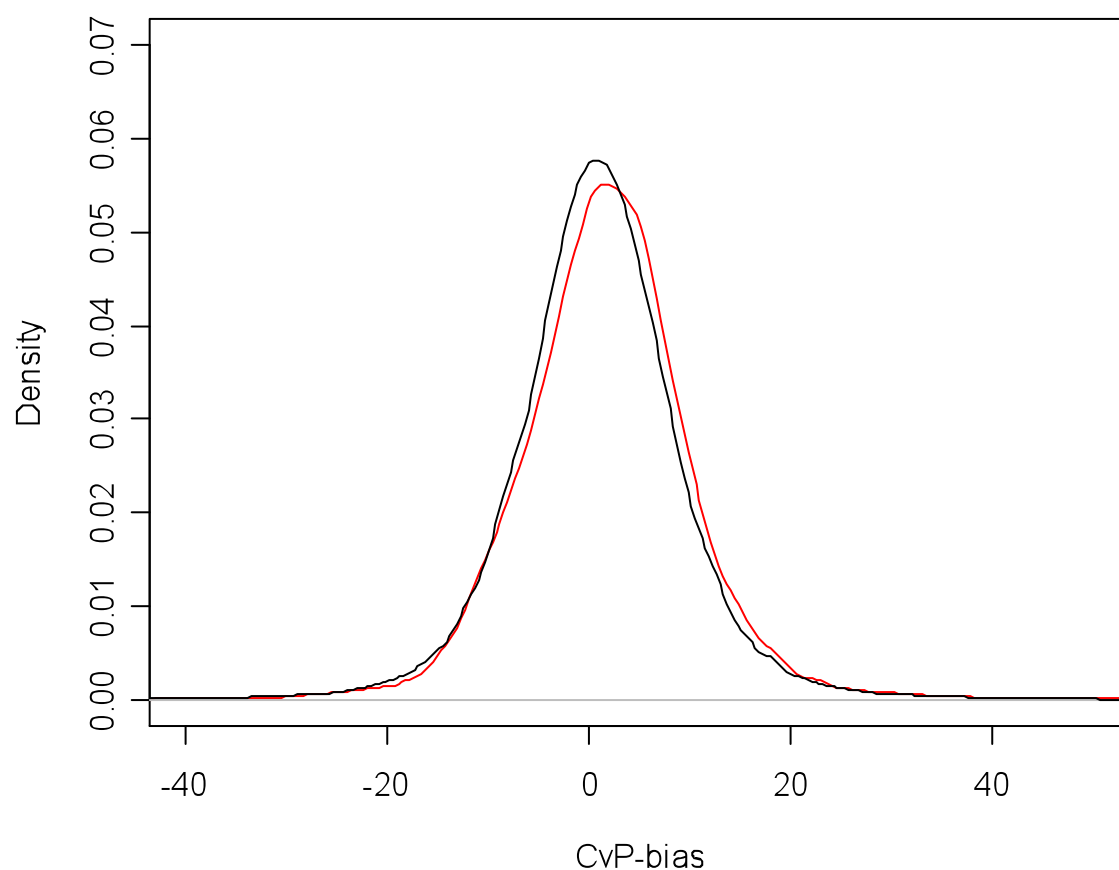


Figure S4. Distribution of an alternative measure of amino acid bias, CvP-bias, across proteins for warm-blooded vertebrates (red line: mammals, birds) and cold-blooded vertebrates (black line: Reptilia, Amphibia, fish). *CvP-bias* is significantly increased in warm-blooded relative to cold-blooded animals (Wilcoxon rank sum test: $p < 10^{-15}$).

Supplemental Tables

Table S1. Summary of the 15 prokaryotic genomes used for the calculation of co-orthologs in this study (see Methods for details). OGT=optimal growth temperature.

Life style	Species	Kingdom	OGT(°C)	Number of orthologs
Hyper-thermophiles	<i>Aquifex aeolicus</i>	Bacteria	90	5334
	<i>Methanopyrus kandleri</i>	Archaea	98	
	<i>Methanococcus jannaschii</i>	Archaea	85	
	<i>Pyrococcus abyssi</i>	Archaea	97	
	<i>Pyrococcus furiosus</i>	Archaea	98	
Thermophiles	<i>Thermoplasma volcanium</i>	Archaea	60	5366
	<i>Thermoplasma acidophilum</i>	Archaea	58	
	<i>Methanobacterium thermoautotrophicum</i>	Archaea	65	
	<i>Carboxydotherrmus hydrogenoformans Z-2901</i>	Bacteria	68	
	<i>Chloroflexus aurantiacus J_10_fl</i>	Bacteria	50-60	
Mesophiles	<i>Haemophilus influenzae</i>	Bacteria	37	4593
	<i>Staphylococcus aureus Mu50</i>	Bacteria	37	
	<i>Corynebacterium glutamicum</i> ATCC 13032	Bacteria	30	
	Bielefeld			
	<i>Mycobacterium tuberculosis H37Rv</i>	Bacteria	37	
	<i>Salmonella typhi</i>	Bacteria	37	

Table S2. *P*-value of two measures of amino acid usage bias, *ERK* and *CvP-bias*, across all codons of orthologous proteins for warm-blooded and cold-blooded animals. *ERK* and *CvP-bias* are significantly higher in warm-blooded genomes compared to *Anolis carolinensis* (1-sided Wilcoxon rank sum tests).

Class	Species	<i>p</i> (<i>ERK</i>)	<i>p</i> (<i>CvP-bias</i>)
Mammalia	<i>Mus musculus</i>	0.0014	0.00045
Mammalia	<i>Rattus norvegicus</i>	0.0093	0.012
Mammalia	Human	4.9×10^{-16}	1.3×10^{-11}
Mammalia	<i>Bos taurus</i>	2.2×10^{-16}	2.2×10^{-16}
Birds	<i>Gallus gallus</i>	2.2×10^{-16}	2.2×10^{-16}
Amphibia	<i>Xenopus laevis</i>	0.081	1
Amphibia	<i>Xenopus tropicalis</i>	0.10	1
Fish	<i>Danio rerio</i>	1	1
Fish	<i>Tetraodon nigroviridis</i>	1	1
Fish	<i>Takifugu rubripes</i>	1	1

Table S3. Comparison of orthologous protein fragments from *Gallus gallus* and three reptilian species *p*-values are from Wilcoxon paired rank tests, comparing mean *ERK* or *CVP-bias* of chicken and the respective reptile.

Reptilia species	Number of protein fragments	Reptile <i>ERK</i>	Chicken <i>ERK</i>	<i>p</i>	Reptile <i>CvP-bias</i>	Chicken <i>CvP-bias</i>	<i>p</i>
Alligator	508	-10.34	-7.41	4×10^{-06}	7.76	11.53	2×10^{-16}
Lizard	138	-13.53	-11.07	0.00027	4.24	7.64	2×10^{-07}
Turtle	429	-11.64	-10.85	0.011	4.91	7.21	7×10^{-12}

Table S4. List of species used in the analysis controlling for phylogenetic non-independence

Species	OGT
<i>Photobacterium profundum</i>	15
<i>Agrobacterium tumefaciens</i> WashU	26
<i>Bradyrhizobium japonicum</i>	26
<i>Fusobacterium nucleatum</i> subsp. <i>nucleatum</i>	26
<i>Mesorhizobium loti</i>	26
<i>Mycoplasma mobile</i> 163K	26
<i>Nitrosomonas europaea</i>	26
<i>Pseudomonas syringae</i> pv. <i>tomato</i>	26
<i>Sinorhizobium meliloti</i>	26
<i>Synechococcus elongatus</i>	26
<i>Xanthomonas campestris</i> pv. <i>campestris</i>	26
<i>Xylella fastidiosa</i>	26
<i>Streptomyces avermitilis</i>	27
<i>Streptomyces coelicolor</i>	27
<i>Oceanobacillus iheyensis</i>	28
<i>Pseudomonas putida</i> KT2440	28
<i>Ralstonia solanacearum</i>	28
<i>Yersinia pestis</i> KIM	28
<i>Borrelia burgdorferi</i>	29
<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i>	29
<i>Bacillus anthracis</i> str. <i>Ames</i>	30
<i>Bacillus halodurans</i>	30
<i>Bdellovibrio bacteriovorus</i>	30
<i>Chromobacterium violaceum</i>	30
<i>Deinococcus radiodurans</i>	30
<i>Desulfovibrio vulgaris</i> subsp. <i>vulgaris</i> str. <i>Hildenborough</i>	30
<i>Lactobacillus plantarum</i>	30
<i>Lactococcus lactis</i> subsp. <i>lactis</i>	30
<i>Leptospira interrogans</i> serovar <i>Copenhageni</i>	30
<i>Rhodopseudomonas palustris</i>	30
<i>Shewanella oneidensis</i>	30
<i>Xanthomonas axonopodis</i> pv. <i>citri</i>	30
<i>Haemophilus ducreyi</i>	32
<i>Corynebacterium glutamicum</i> ATCC 13032	33
<i>Methanococcus maripaludis</i>	35
<i>Rickettsia conorii</i>	35
<i>Rickettsia prowazekii</i>	35
<i>Bordetella bronchiseptica</i>	36
<i>Methanosarcina mazei</i>	36
<i>Bacillus subtilis</i>	37
<i>Bacteroides thetaiotaomicron</i>	37
<i>Bifidobacterium longum</i>	37
<i>Bordetella parapertussis</i>	37
<i>Bordetella pertussis</i>	37
<i>Brucella melitensis</i>	37

<i>Chlamydia muridarum</i>	37
<i>Chlamydia trachomatis</i>	37
<i>Chlamydophila pneumoniae</i> AR39	37
<i>Clostridium acetobutylicum</i>	37
<i>Clostridium perfringens</i>	37
<i>Clostridium tetani</i>	37
<i>Corynebacterium diphtheriae</i>	37
<i>Corynebacterium efficiens</i>	37
<i>Enterococcus faecalis</i>	37
<i>Escherichia coli</i>	37
<i>Haemophilus influenzae</i>	37
<i>Halobacterium</i> sp. NRC-1	37
<i>Helicobacter hepaticus</i>	37
<i>Helicobacter pylori</i>	37
<i>Lactobacillus johnsonii</i>	37
<i>Listeria innocua</i>	37
<i>Listeria monocytogenes</i>	37
<i>Mycobacterium bovis</i>	37
<i>Mycobacterium leprae</i>	37
<i>Mycobacterium tuberculosis</i> CDC1551	37
<i>Mycoplasma gallisepticum</i>	37
<i>Mycoplasma genitalium</i>	37
<i>Mycoplasma penetrans</i>	37
<i>Mycoplasma pneumoniae</i>	37
<i>Mycoplasma pulmonis</i>	37
<i>Pasteurella multocida</i>	37
<i>Porphyromonas gingivalis</i>	37
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi Ty2	37
<i>Salmonella typhi</i>	37
<i>Shigella flexneri</i>	37
<i>Shigella flexneri</i> 2a str. 2457T	37
<i>Staphylococcus epidermidis</i>	37
<i>Streptococcus agalactiae</i> serogroup III	37
<i>Streptococcus mutans</i>	37
<i>Streptococcus pneumoniae</i> R6	37
<i>Streptococcus pyogenes</i>	37
<i>Treponema denticola</i>	37
<i>Treponema pallidum</i>	37
<i>Vibrio cholerae</i>	37
<i>Vibrio parahaemolyticus</i>	37
<i>Vibrio vulnificus</i>	37
<i>Wolinella succinogenes</i>	37
<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i>	39
<i>Campylobacter jejuni</i>	40
<i>Methanosarcina acetivorans</i>	40
<i>Pseudomonas aeruginosa</i>	40
<i>Chlorobaculum tepidum</i>	48

<i>Thermoplasma acidophilum</i>	59
<i>Thermoplasma volcanium</i>	60
<i>Thermoanaerobacter tengcongensis</i>	75
<i>Sulfolobus solfataricus</i>	80
<i>Sulfolobus tokodaii</i>	80
<i>Thermotoga maritima</i>	80
<i>Nanoarchaeum equitans</i>	82
<i>Archaeoglobus fulgidus</i>	83
<i>Aquifex aeolicus</i>	85
<i>Methanocaldococcus jannaschii</i>	85
<i>Thermus thermophilus</i> HB27	85
<i>Aeropyrum pernix</i>	95
<i>Pyrococcus abyssi</i>	96
<i>Methanopyrus kandleri</i>	98
<i>Pyrococcus horikoshii</i>	98
<i>Pyrobaculum aerophilum</i>	100
<i>Pyrococcus furiosus</i>	100
