Supporting Information

Johnson and Wang 10.1073/pnas.1000704107

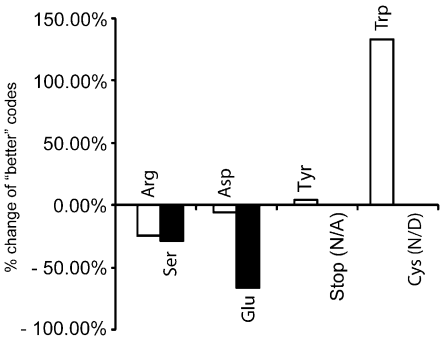

Fig. S1. Codon block expansion analysis of the other four split blocks. There was no difference between Arg and Ser after expansion. The Asp-Glu codon block appeared to show a similar phenomenon as those amino acids in Fig. 3, as Glu had a large reduction in the number of better codes. However, Glu had an extremely low anticodon enrichment value (0.596) (Fig. 1*C*), which increased to 0.837 upon Glu expansion to occupy the entire GAN block. A <1 enrichment value suggests no significant enrichment of anticodons near Glu in the ribosomal structures. Therefore, we concluded that the finding described in Fig. 3 was not applicable to the Asp-Glu codon block. Expansion of the stop signal was not applicable because of the nonexistence of a stop entity in proteins, and Cys expansion could not be determined because of its too few presence at ribosomal protein surface.

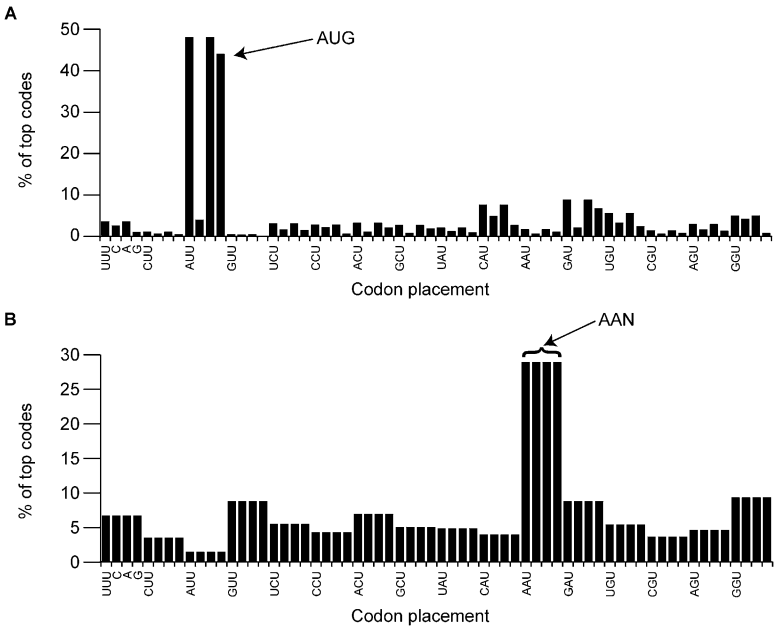

Fig. S2. Analysis of the top 1,000 random codes reveals evidence for codon reassignment during code evolution. The probability of finding an amino acid at a specific location in the 1,000 random codes with the highest global average enrichment values was determined (see the *Methods* for how Ile and Pro are placed in the random codes). (*A*) The placement of Ile in the top 1,000 codes reveals a strong preference for the methionine codon AUG because of the strong enrichment of Ile near the corresponding anticodon CAU in the ribosome. (*B*) The placement of Pro in the top 1,000 codes reveals a strong preference for the AAN codon block due to the ribosomal enrichment of Pro near the UU moiety.

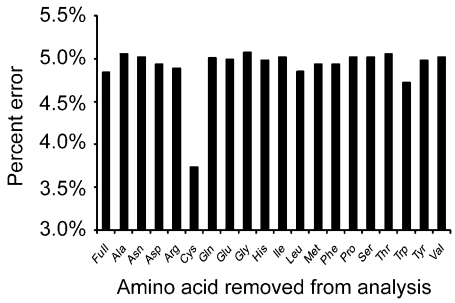

Fig. S3. Inclusion of Cys in the global analysis dramatically increases the deviation of the randomized genetic codes. Each amino acid was excluded from global analysis of ribosomal anticodon-amino acid enrichment individually and the percent error of the average enrichment values from 10,000 codes was calculated. Only the removal of Cys shows a marked effect on the percent error in the populations.

Table S1. Results for correlation analysis of ribosomal anticodon-amino acid enrichment and the canonical code are consistent regardless of which random code generator was used

Method	Specific subset	Optimal subset
RAND	99.9555%	95.1925%
NNY	99.9971%	95.7479%
SYN	99.9996%	97.8008%

One million random codes were generated from each generator and Monte Carlo analysis was performed using both the specific and optimal subset of amino acids. The percentage of codes with a lower average enrichment than that of the canonical code is shown.