**Supplementary Online Materials:**

Materials and Methods
Figures S1-S15
Tables S1-S7
References

# Supplementary Online Materials:

## Materials and Methods:

### Cell Culture

The following cell lines were used for this study: Epstein-Barr virus (EBV)-transformed lymphoblastoid cell lines (LCLs) derived from 5 human (Coriell YRI, NIGMS Human Genetic Cell Repository, GM18505, GM18507, GM18516, GM19193, GM19204), and 5 chimpanzee (*Pan troglodytes*) individuals (New Iberia Research Center: Min 18358, Min 18359; Coriell/IPBIR: NS03659, NS04973, Arizona State University, Pt91), and rhesus *Herpesvirus papio* transformed LCLs from 5 rhesus macaque (*Macaca mulatta*) individuals (Harvard Medical School, NEPRC: 150-99, R181-96, R249-97, 265-95, R290-96). Cells were maintained at identical conditions of 37° with 5% $CO_2$ in RPMI media with 15% FBS, supplemented with 2 mM L-glutamate, 100 IU/ml penicillin, and 100 µg/ml streptomycin. Internal standard LCL (Coriell YRI, NIGMS Human Genetic Cell Repository, GM19238) was grown in RPMI minus L-Lysine and L-Arginine, 15% dialyzed FBS, and L-$^{13}C_6{}^{15}N_4$-arginine (Arg-10) and L-$^{13}C_6{}^{15}N_2$-lysine (Lys-8) (Cambridge Isotopes, Andover, MA, USA) supplemented with 2 mM L-glutamate, 100 IU/ml penicillin, and 100 µg/ml streptomycin under identical conditions as the unlabeled LCLs. The internal standard LCL was grown for 6 doublings to assure complete SILAC label incorporation. Complete label incorporation was verified by analyzing the protein lysate from the labeled LCL alone by high-resolution LC-MS/MS.

### mRNA and Protein Quantification Strategy Overview

We used a protein quantification strategy where a single human SILAC labeled LCL GM19238 served as an internal standard for comparing protein levels across individuals and species (fig. S1). We chose SILAC because it does not suffer from the dynamic range compression and reduced quantitative accuracy issues of MS/MS quantification techniques such as multiplexed iTraq labeling, nor does it suffer from the high level of noise present in label-free quantification techniques (*21*). We also collected RNA-seq data for the same cell line used as an internal standard for the SILAC labeling allowing us to derive $\log_2$(sample/standard) ratios that were comparable across the two technologies (fig. 1A-1C). Thus, we could use the same statistical procedures for testing for mRNA as well as protein expression level differences.

### Quantitative, High-Resolution Mass Spectrometry

LCLs were washed in PBS three times and then lysed using the UPX Universal Protein Extraction Kit (Expedeon Inc., San Diego, CA, USA). Protein quantitation was performed using the Qubit fluorometry assay (Invitrogen, Carlsbad California, USA) and the reducing agent-compatible (RAC) version of the BCA Protein Assay (Thermo Scientific, Waltham, Massachusetts, USA). 12µg of each sample was combined with 12µg of the SILAC labeled lysate from human LCL GM19238. Note that the SILAC lysate was prepared once and used as an internal standard through the quantification of the 15 cell lines. 24µg of each combined sample was then processed by SDS-PAGE using a 4-12% Bis Tris NuPage mini-gel (Invitrogen, Carlsbad California, USA). Calibration was with Thermo PageRuler broad range markers. Each of 40 gel segments were processed by in-gel digestion using a ProGest robot (DigiLab, Marlborough, MA, USA) with the following protocol: wash with 25mM ammonium bicarbonate followed by acetonitrile, reduce with 10mM dithiothreitol at 60°C followed by alkylation with 50mM iodoacetamide at room temperature, digest with trypsin (Promega, Madison, WI, USA) at 37°C for 4h, and quench with formic acid. The supernatant was analyzed directly without further processing. Each gel digest was analyzed by nano-LC/MS/MS with a Waters NanoAcquity HPLC system interfaced to a ThermoFisher LTQ-Orbitrap Velos Pro. Peptides were loaded on a trapping column and eluted over a 75µm analytical column at 350nL/min using a 1-hour LC gradient. Both columns were packed with Jupiter Proteo resin (Phenomenex, Torrance, California, USA). The mass spectrometer was operated in data-dependent mode, with MS performed in the Orbitrap at 60,000 FWHM resolution and MS/MS performed in the LTQ. The fifteen most abundant ions were selected for MS/MS.

### Computational Analysis of Mass Spectra

Low-level analysis was performed using the open-source proteomics software tool PVIEW (Release December 23, 2012; http://compbio.cs.princeton.edu/pview)(*22*). As input to PVIEW, we generated *in silico* translations of coding genes from the UCSC Genome Browser database based on gene models from build hg19 of the human genome. Each protein sequence entry retained the corresponding Ensembl gene identifier and gene symbol. Database searches were performed using ±4 p.p.m. MS1 tolerance and an MS2 window tolerance of ±0.5 Da. Up to 2 missed tryptic

cleavages were allowed during search. Carboxyamidomethylation of cysteine was used as fixed modification. Up to two methionine oxidations were allowed as variable modifications of a tryptic peptide. Peptide spectrum matches were obtained at a stringent false discovery rate (FDR) of 1%. We used the median $\log_2$(sample/standard) ratio across all independent quantifications of a protein (distinct peptides including duplicate peptide measurements across fractions and for differing charge states). Note that use of the hg19 database was sufficient as SILAC pairs as the expected isotope shift used for quantification were only present for peptides that that had the same underlying sequence in the human internal standard line (fig. S1). The $\log_2$(sample/standard) quantifications were normalized per cell line by shifting them so that their median was centered at zero, thus, accounting for small differences in input protein sample

**RNA-seq Analysis**

RNA-seq libraries were prepared using the TruSeq RNA Sample Prep Kits (Illumina, San Diego, CA, USA). Prior to library preparation, we confirmed that the RNA was of high quality using Agillent's Bioanalyzer 2100 (Santa Clara, CA, USA). 50bp-long single end reads were then sequenced on an Illumina Hi-Seq 2000. 27 RNA-seq libraries were sequenced across 6 lanes of a flow cell (3 sets of 9 samples with unique bar codes were each sequenced across 2 lanes). The present study only includes the RNA-seq samples that had matching protein data. Reads were aligned using BWA with the following command-line parameters "bwa aln -n 2 -t 3" and "bwa samse -n 1" Alignments with quality score < 10 were filtered using "samtools view -S -h -q 10 –b." Reads were aligned to genome builds hg19 (human), Pantro3 (chimpanzee), and RheMac2 (rhesus) for each respective species. We used a curated set of orthologous exons (http://giladlab.uchicago.edu/orthoExon) to derive expression measurements for each gene(*14*). Reads overlapping a single orthologous exons in each species were counted to derive reads per kilo-base per million mapped reads (RPKM) values for a gene in each species per cell line. Because we also had RNA-seq data for the cell line used as the internal standard in the SILAC experiment (GM19238), we were able to compute a $\log_2$(sample/standard) ratios of the RPKM values obtained for each line relative to the internal standard analogous to the SILAC experiment. The ratios were normalized so that their median was centered at zero to account for minor differences in total number of sequencing reads from the analyzed cell line and the standard. Note, in contrast to LC-MS/MS, RNA-seq allows higher depth of coverage of the genome. Using the RPKM values from the internal standard cell line, we found that medium to highly expressed genes were measured at the protein level by our protein quantification strategy in all 3 species in 3 or more individuals (fig. S16).

**Differential Expression Analysis of mRNA and Protein Measurements**

Computational analysis of SILAC mass spectra and the RNA-seq analysis of reads relative to an internal standard cell line allowed us to measure 3,390 genes in 3 or more individuals in human, chimpanzee, and rhesus macaque. Thus, for each gene and for each species we had at most five $\log_2$(sample/standard) RPKM mRNA and five $\log_2$(sample/standard) SILAC protein measurements, a total of 15 mRNA and 15 protein measurements per gene (table S4). We used likelihood ratio (LR) tests within the framework of nested linear models to compare the fit of null models that correspond to patterns of no expression difference between species, to models that correspond to patterns of gene expression differences between pairs of species (human-chimpanzee, human-rhesus, and chimpanzee-rhesus). We used a similar framework to analyze both the mRNA and protein expression data, and obtained a total of 6 p-values per gene and 3,390 p-values per pairwise comparison. Specifically, to test the null of no expression difference and obtain a p-value, we compared the likelihood ratio of a model of *expression = c + epsilon* (2 degrees of freedom) to the model *expression = species * beta + c + epsilon* (3 degrees of freedom) after they were fit to the data from the pairwise species comparison separately for either the mRNA or protein expression data. Two times the difference between the log-likelihoods follows a $\chi^2$ distribution with 1 degree of freedom from which we obtained the p-value. To control for multiple testing across a pairwise species comparison, we estimated the false discovery rate (FDR) using QVALUE (*23*). Proteins or transcripts were classified as DE at an FDR of 1% by applying a threshold of 0.01 to the computed q-value for the comparison. In total, each gene could be assigned 6 indicators 0=not-DE or 1=DE for each pairwise comparison for both mRNA and protein (table S4).

**Comparison of mRNA and Protein Divergence**

When we compare the results of inter-species DE at the mRNA and protein levels, a discrepancy in classification of DE could reflect cases of DE at only one level (e.g., mRNA only) or reflect cases where there is a higher level of within-species variation in one level only (e.g., higher between-individual variation in protein levels). To distinguish between these possibilities we compared estimates of pairwise species divergence at the mRNA level to divergence at the protein level. Focusing on the comparison between human and chimpanzee, we randomly paired 5 human individuals with 5 chimpanzee individuals. We then computed a $\log_2$(human/chimpanzee) ratios using mRNA

and protein levels, normalizing the contribution of the internal standard cell line (Fig. 1D). In order to determine if the divergence estimated using mRNA levels differed from divergence estimates obtained from protein levels, we used a likelihood ratio (LR) test. We compared the likelihood ratio of the model *divergence* = *c* + *epsilon* (2 degrees of freedom) to *divergence* = *technology* * *beta* + *c* + *epsilon* (3 degrees of freedom). Here, the indicator *technology* is mRNA or protein. Two times the difference between the log-likelihoods follows a $\chi^2$ distribution with 1 degree of freedom from which we obtained p-value for the difference between these divergence estimates. Using this test, we obtained 3,390 p-values comparing divergence estimates obtained using mRNA measurements to those obtained using protein measurements. To control for multiple testing, we estimated false discovery rate (FDR) using QVALUE (*23*). We applied a threshold of 0.01 to q-values to identify genes where mRNA divergence differed significantly from protein divergence at an FDR of 1%.

**Identification of Patterns Consistent with Stabilizing Selection**

A pattern of low variation in expression levels, both within and between species, is consistent with the action of stabilizing selection on gene regulation. To identify such patterns, we first excluded genes that are differentially expressed between species (using a likelihood ratio test statistic assuming a $\chi^2$ distribution at an FDR of 1%). We then ranked the remaining genes based on overall variation in gene expression levels across individuals and species (from low to high variance). The regulation of genes at the top of the ranked list are more likely to have evolved under stabilizing selection than genes at the bottom of the list. Explanations based on environmental effects for patterns of low variation within and between species are unlikely, as differences in environment (both within and between species) will tend to increase the observed variance. The regulation of genes with constant mRNA or protein expression levels, across individuals and species, is therefore likely to be genetically controlled. The exact rank below which the expression patterns do not provide compelling evidence for stabilizing selection can only be chosen arbitrarily. In the main paper we used the cutoff of the top 300 genes – but importantly, we confirmed that our conclusions were robust with respect to a large range of cutoffs.

**Identification of Patterns Consistent with Lineage-Specific Directional Selection**

A pattern of a lineage-specific shift in expression level, accompanied by low variation in expression levels within species, is consistent with the action of directional selection on gene regulation. To identify patterns that are consistent with human lineage-specific directional selection, we used two steps. First, we excluded genes that are differentially expressed between chimpanzee and rhesus (at 1% FDR). We then assumed that, for the remaining genes, the chimpanzee and rhesus individuals had the same mean expression level, and identified, using a likelihood ratio statistic, genes that are differentially expressed between humans and the non-human primates (at FDR 1%). We used a similar procedure to identify expression patterns consistent with chimpanzee lineage-specific directional selection. When the lineage-specific change in expression was accompanied by an increase in within-species variance in expression levels (at an F-test $P < 0.05$), we inferred a lineage-specific relaxation of constraint. As previously discussed(*4*), the patterns we identify carry a significant confounding caveat, they might be a consequence of a specific environmental influence on that lineage rather than selection of a genetic change.

**Gene Ontology Analysis**

Gene Ontology enrichment analysis was performed using DAVID's functional enrichment analysis tool (http://david.abcc.ncifcrf.gov)(*24*). We searched for enrichments using the GOTERM_BP_FAT, GOTERM_CC_FAT, GOTERM_MF_FAT, and KEGG_PATHWAY annotation categories. As a background list we used the 3390 genes quantified at the protein level in least 3 individuals from each of the three species. The count threshold was kept at the default 2 genes. We report the modified Fisher Exact P-value (EASE score) computed by the DAVID system.

**$d_N$ and $d_S$ Analysis**

$d_N$ and $d_S$ values were obtained from one-to-one cross species orthologs from the Ensembl database. We excluded cases where synonymous divergence was 0, and set $d_N/d_S$ ratio to 0 when non-synonymous divergence was 0.

**Analysis of Gene Features**

We used the median length value for gene features (5'UTR, coding sequence, transcript, and 3'UTR) for all known human isoforms of a gene from human genome build hg19. We used the August 2010 release (http://www.microrna.org) of predicted micro-RNA binding sites from miRanda to test whether genes consistent with buffering or compensatory selection had increased or decreased numbers of predicted micro-RNA binding sites. We found no significant enrichment of micro-RNA binding sites in this set of genes.

**Tissue-Specific Gene Expression Analysis**

We used the cross-tissue gene expression data set from BioGPS (http://biogps.org/downloads) (*25*). We only used the probes from the Affymetrix U133a array ignoring data from the custom GNF1H array probes. The data set does not report transcripts as present/absent, but as median normalized MAS5 algorithm intensities. To address this, we used a global $\log_2$-fold cutoff, that is, a gene was considered expressed in that tissue if the MAS5 intensity was above our cutoff. We confirmed that our results hold for a large range of $\log_2$-fold intensity cutoffs as low as 1.5 and as high as 8.5.

**Protein-Protein Interaction Analysis**

Human physical protein-protein interactions were obtained from BIOGRID v.3.2.98. (http://thebiogrid.org). We removed any redundant interactions from the set. Interaction counts were merged with our list of genes based on gene symbol.

**PTM Analysis**

Known phosphorylation and ubiqutination site data was obtained from PhosphositePlus (http://www.phosphosite.org; last modification date of "Mon Mar 04 09:31:05 EST 2013")(*26*).

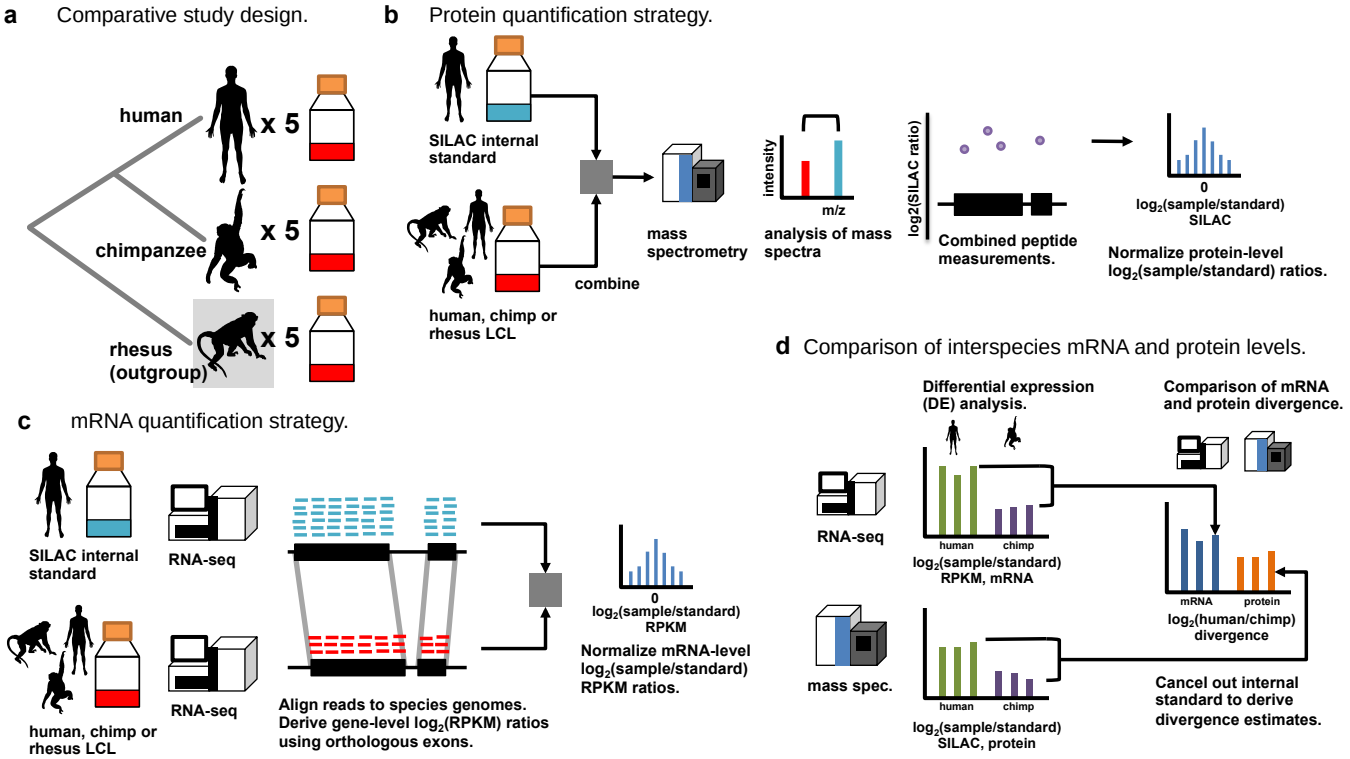


**Figure S1.** Comparative transcript and protein expression level profiling. (a) mRNA and protein levels were compared in lymphoblastoid cell lines (LCLs) derived from 5 human, 5 chimpanzee, and 5 rhesus individuals. (b) Proteins were quantified using a stable isotope-labeled LCL, which served as a common standard. Protein expression levels were estimated as $\log_2$(sample/standard) SILAC ratio across quantified peptides. The median of all protein-level quantifications was shifted to zero to correct for small differences in input protein between the analyzed cell line and internal standard. (c) mRNA levels were quantified by sequencing RNA from all LCLs (including the common standard line). Reads were aligned to species specific genomes and reads per kb of exon per million mapped reads (RPKM) ratios between the analyzed cell line and the internal standard LCL were computed. To correct for minor differences in sequencing read depth, the median $\log_2$(sample/standard) RPKM ratio across measured transcripts was centered at zero. (d) An illustration of the inter-species comparisons performed for a single gene measured in at least 3 humans and 3 chimpanzees. Individuals across species were paired to derive measurements of mRNA and protein $\log_2$(human/chimpanzee) divergence to detect genes where these estimates differed significantly.

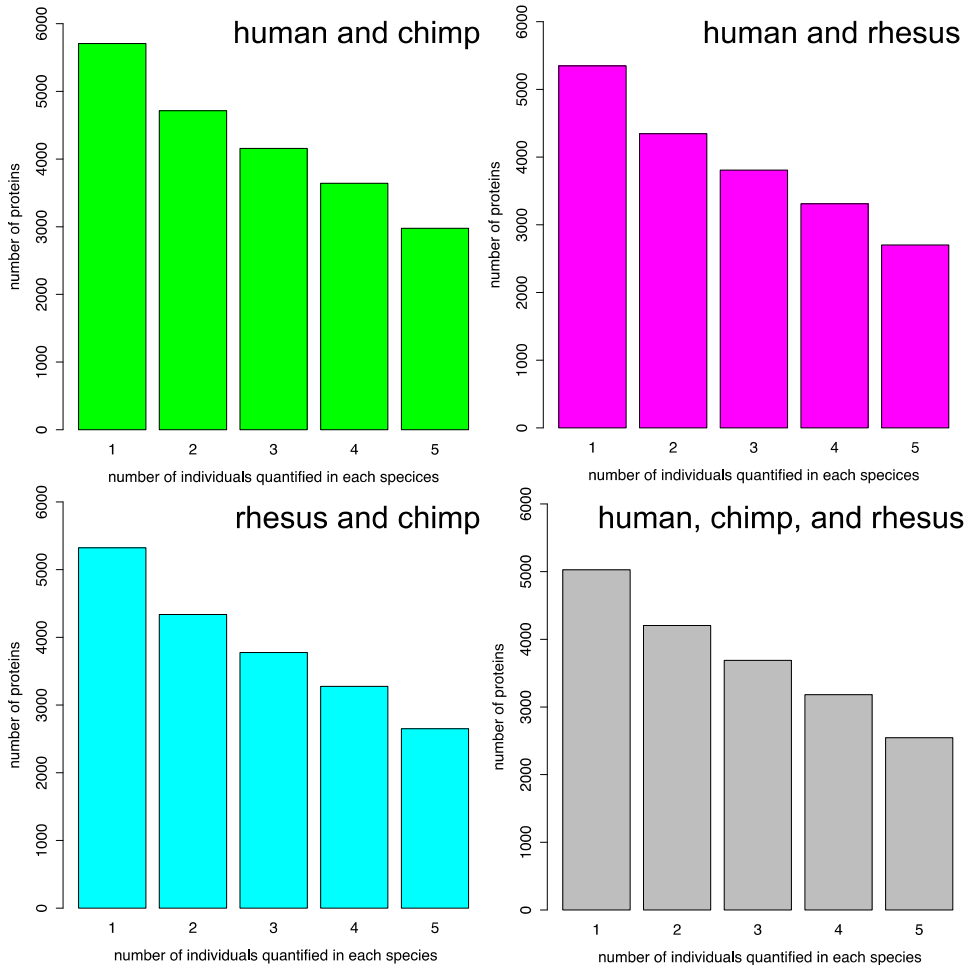


**Figure S2. Reproducibility of Protein Quantifications.** Histograms of the number of proteins quantified in up to 5 individuals in each pairwise species comparison (green, magenta, and cyan), and histogram of the number of proteins quantified in all 3 species in up to 5 individuals.

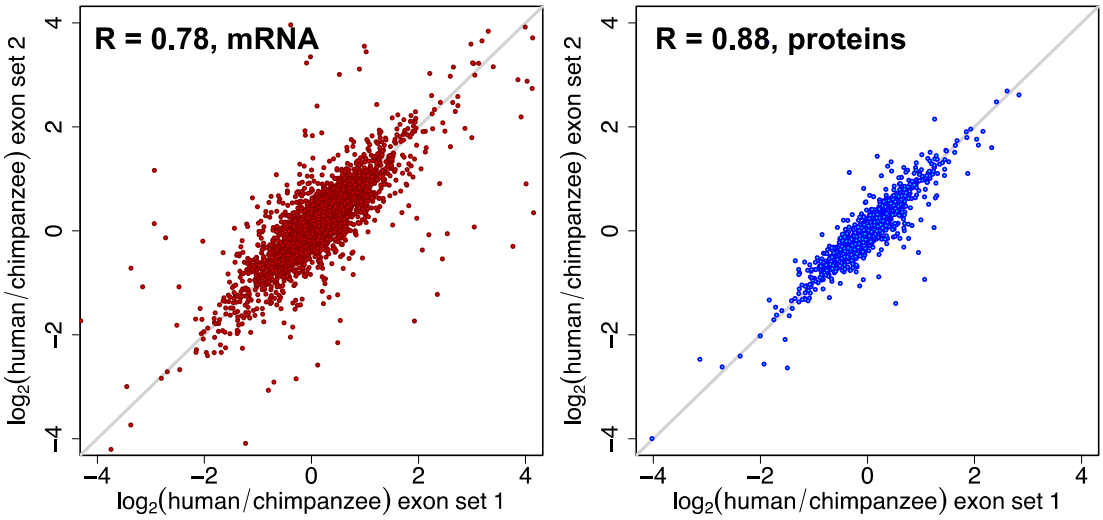


**Figure S3. Comparison of the consistency of mRNA and protein measurements across orthologous exons.** We estimated technical variation of our mRNA and protein divergence estimates by computing divergence using the median RPKM values and SILAC peptide ratios from two randomly generated and non-overlapping subsets of orthologous exons from the same gene (exon set 1 and exon set 2 in the scatterplots above). Protein measurements were obtained by mapping peptide sequences back to orthologous exons. Each point corresponds to the median divergence estimate obtained from exons in each respective set. R-values correspond to Pearson's correlation. Our analysis indicates that our observation of buffering or compensation at the protein level cannot be explained by differences in technical variation between mRNA and protein measurements.

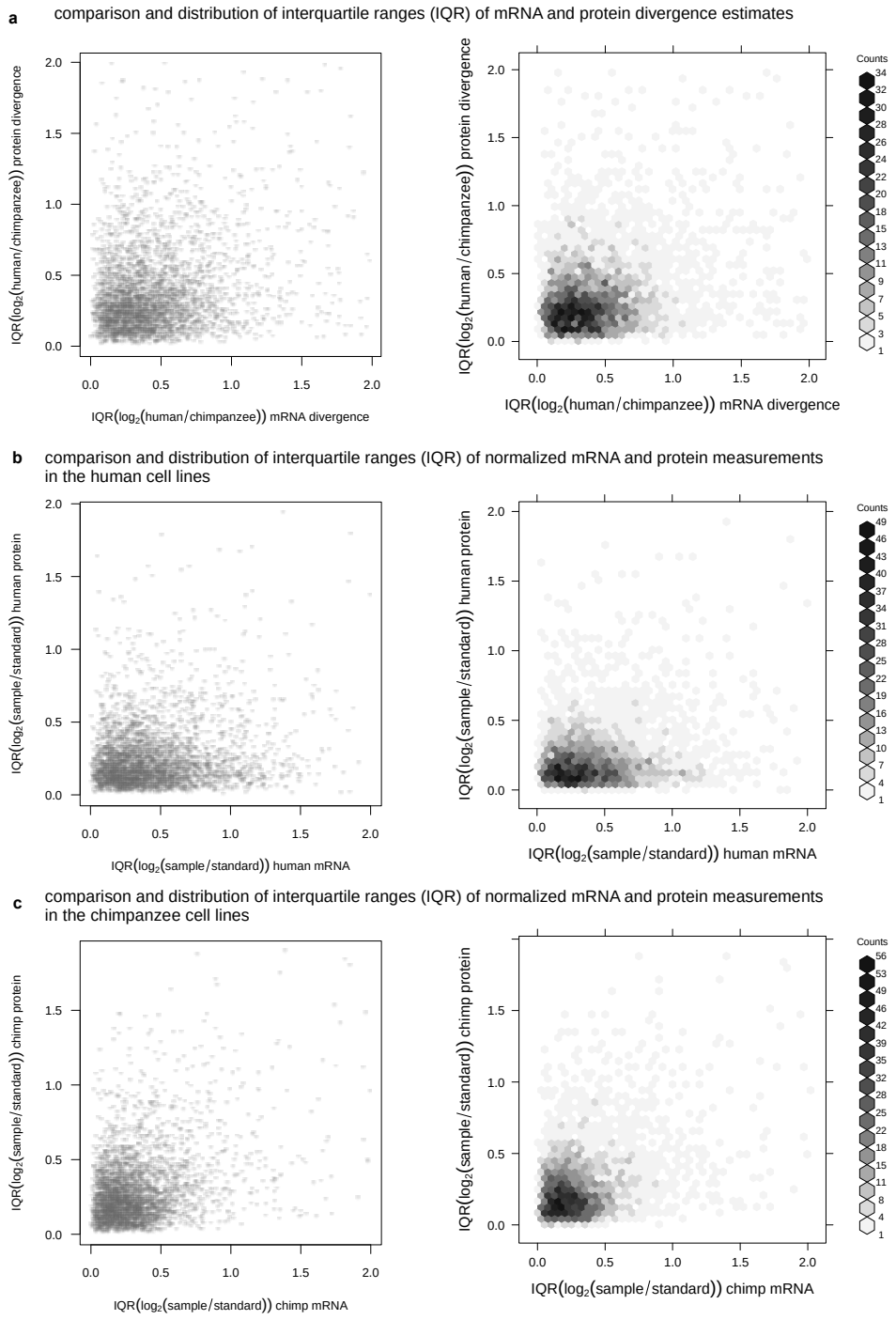


**Figure S4. Comparison of biological variation in mRNA and protein measurements.** (a) Left, scatterplot of the interquartile ranges (IQRs) for mRNA (x-axis) and protein divergence (y-axis), that is $\log_2$(human/chimpanzee), for the 3,390 genes analyzed in this study across LCLs derived from 5 human and chimpanzee individuals. Right, a hexagonal binning plot illustrating the number of data points present in the same region of the scatter plot to the left. Scatter and hexagonal binning plots of IQRs for $\log_2$(sample/standard) RPKM mRNA ratios (x-axis) and $\log_2$(sample/standard) SILAC protein ratios (y-axis) for the (b) human and (c) chimpanzee LCLs. Interquartile ranges were computed using the IQR() function in the R language which uses the following rules for the sample sizes in our study: for measurements in 3 individuals the IQR was set to the maximum and minimum values; for measurements in 4 individuals the IQR was set to half the difference between the 2$^{nd}$ and 3$^{rd}$ ranked values plus half the difference between the 3$^{rd}$ and 4$^{th}$ ranked values. Our results illustrate that the spread of the IQRs obtained using protein measurements were not significantly larger than that of mRNA measurements suggesting our results could not be explained by higher biological variation the protein level.

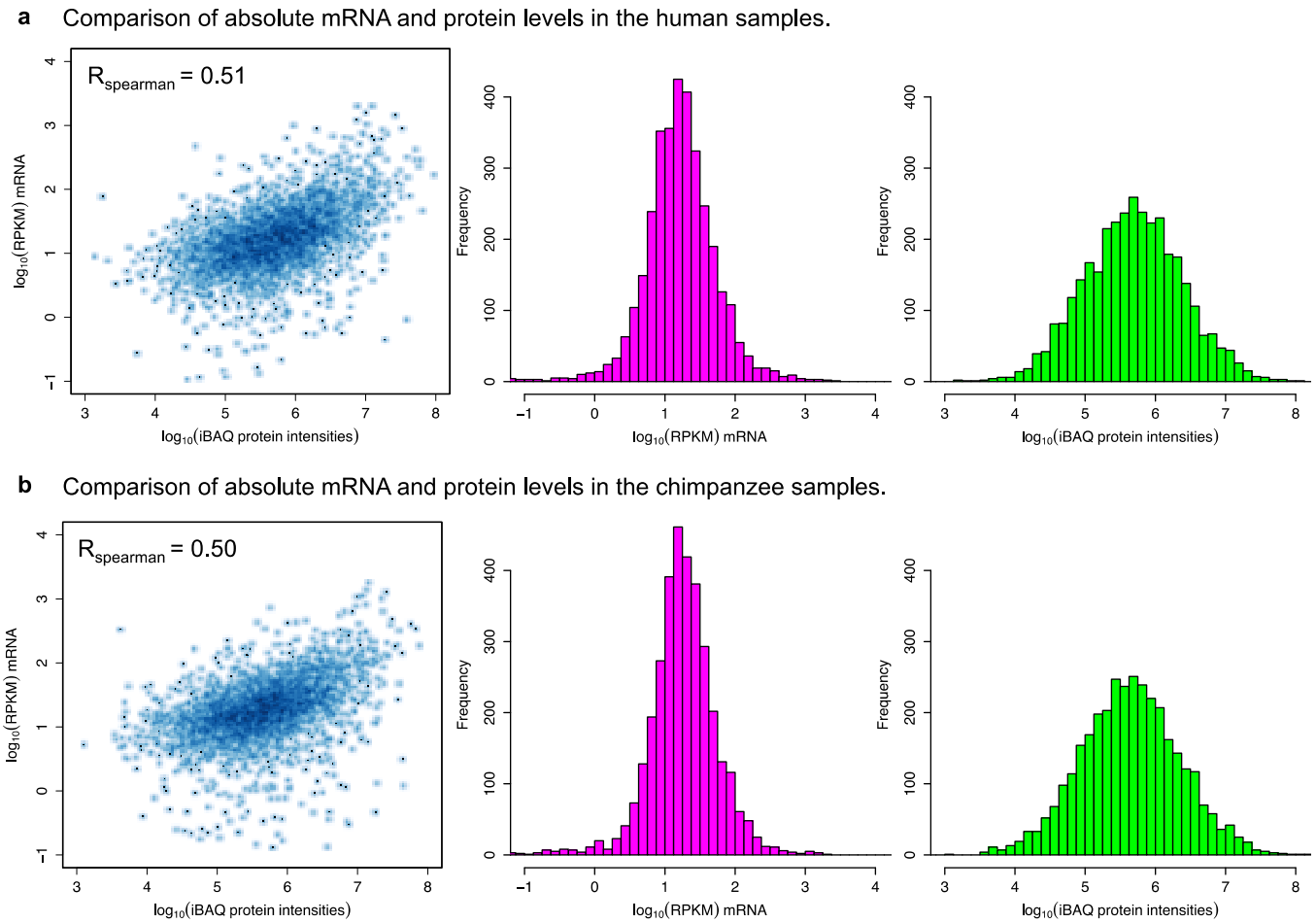


**Figure S5. Comparison of the dynamic range of absolute mRNA and absolute protein levels.** We estimated absolute mRNA and absolute protein expression levels from our data for the 3,390 genes analyzed. In particular, absolute mRNA levels were estimated using RPKM values obtained by combining read counts from each of the 5 individuals across orthologous exons. Absolute protein levels were estimated using iBAQ protein intensities (*27*). We computed iBAQ protein intensities by summing the chromatographic intensities from unlabeled peptides detected across each of 5 individuals for each species above and dividing by the number of tryptic peptides that could be obtained from the protein in the 6-40 amino acid range, thus providing a measure of absolute protein level. We note that iBAQ intensities have proven quite useful for analyzing global trends in absolute protein levels. Yet, they are not as accurate as SILAC for analyzing and detecting small differences between individual protein levels. Thus, any individual iBAQ absolute protein level must be interpreted with caution without additional confirmation from a method like SILAC. Scatter plots to the left show the distribution of RPKM and iBAQ protein intensities. The plots have been adjusted to have the same aspect ratio. Middle and right plots provide histograms of the mRNA RPKM (magenta) and protein iBAQ (green) absolute level estimates, respectively. Histogram bin widths are the same in both plots. mRNA measurements span, approximately, 3 orders of magnitude while absolute protein levels span approximately 4 orders of magnitude for both (a) human and (b) chimpanzee. This result is consistent with previous work and indicates our observation of buffering at the protein expression level is conservative with respect to the differences in dynamic range between mRNA and protein expression measurements.

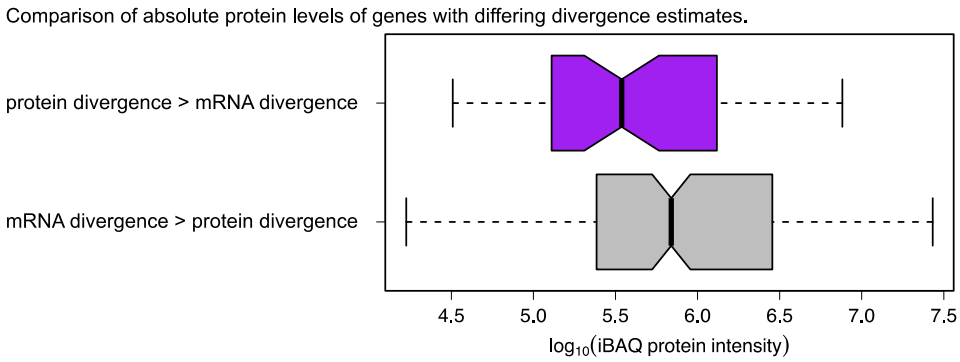


**Figure S6. Stability of protein levels is unlikely due to background quantifications**. Proteins that appear stable between species could be enriched in low, background quantifications. These proteins may appear invariant because their absolute levels are close enough to background that divergence cannot be accurately measured by SILAC. Above are boxplots illustrating the distributions of absolute protein levels estimated using iBAQ protein intensities (see previous figure for details on computation of iBAQ protein intensities) for genes whose protein levels have diverged faster than mRNA levels (purple) and for genes where protein levels are stable relative to mRNA levels (gray) between species (two classes of genes in Fig. 1C. and Fig. 1D). Interestingly, genes with diverged protein levels have lower median iBAQ protein intensities than genes where protein levels are relatively stable to mRNA levels ($p < 0.02$, Wilcoxon rank sum test). This is the opposite of what one would expect if stable protein levels were enriched with background quantifications. In other words, we are able to detect higher protein divergence than mRNA divergence at lower absolute protein levels. Thus, our results are unlikely to be explained by insufficient, background quantifications of proteins that are stable between species.

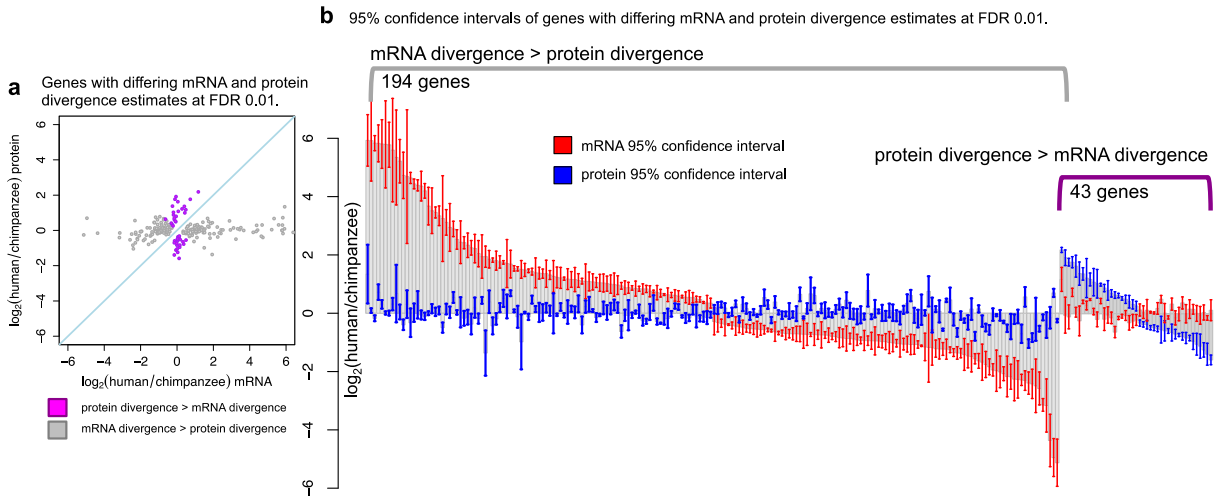


**Figure S7. Analysis of protein divergence estimates obtained by computing the median exon expression measurement**. Throughout the paper, SILAC protein levels were derived using the median $\log_2$(sample/standard) measurement of all peptides measured from a protein. In contrast, mRNA $\log_2$(sample/standard) RPKM ratios were derived from read counts across orthologous exons. We asked if our results changed if we first computed an orthologous exon-specific protein $\log_2$(sample/standard) ratio and then combined measurements by computing the median across the exons. We find that our results are robust with respect to this change. Consistent with the results reported in Fig. 1C and Fig. 1D, among genes whose inter-species mRNA and protein divergence differ, we observed higher absolute divergence at the mRNA level for more genes than that at the protein level.

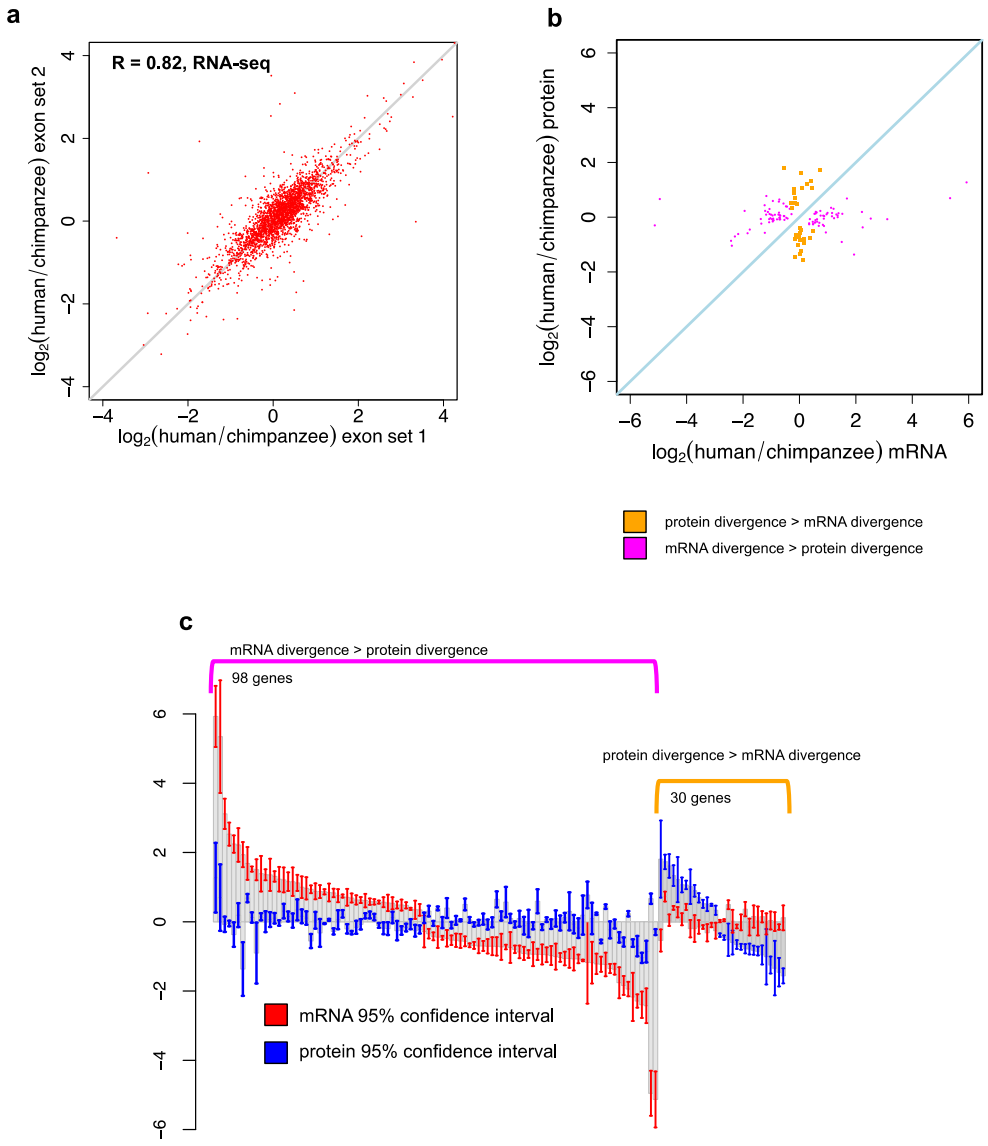


**Figure S8. Analysis of the subset of genes where the RNA-seq read coverages across orthologous exons are significantly correlated (Kendall's τ > 0.4).** The pattern of higher mRNA vs. protein divergence could potentially be explained by a high rate of false positives among transcripts classified as differentially expressed between species. These possible false positives can have a biological explanation, such as intron exclusion or inclusion specific to one species, or technical explanation, such as incorrect assignment of orthologous exons. To exclude these possible explanations, we applied a stringent filtering step and considered only genes where the pattern of read depth was similar across all orthologous exons from that gene. To do so, we computed average normalized read depth in 15-bp windows across an orthologous exon in human and in chimpanzee. We then computed the Kendall's τ correlating the coverage values at the same position in human and in chimpanzee orthologous exons and removed genes where the correlation across all orthologous exons was less than or equal to 0.4. We note that such an approach may also filter out large, true positive mRNA expression differences between species and requires coverage to be sufficiently high in both species for a high correlation to be observed. In total this filtering left 2,582 genes (out of 3,390). As illustrated in (a) the R-value of divergence estimates obtained from differing exons from the same gene improves from 0.78 (see fig. S3) to 0.82 in (a) and the standard deviation of the difference between exons decreases from 0.62 to 0.45 indicating that technical noise is removed by this test. (b) and (c) show that our results are consistent with Fig. 1C and Fig. 1D. In this subset of genes, among those with differing mRNA and protein divergence estimates, we observe more genes with higher absolute mRNA divergence than genes with higher absolute protein divergence.

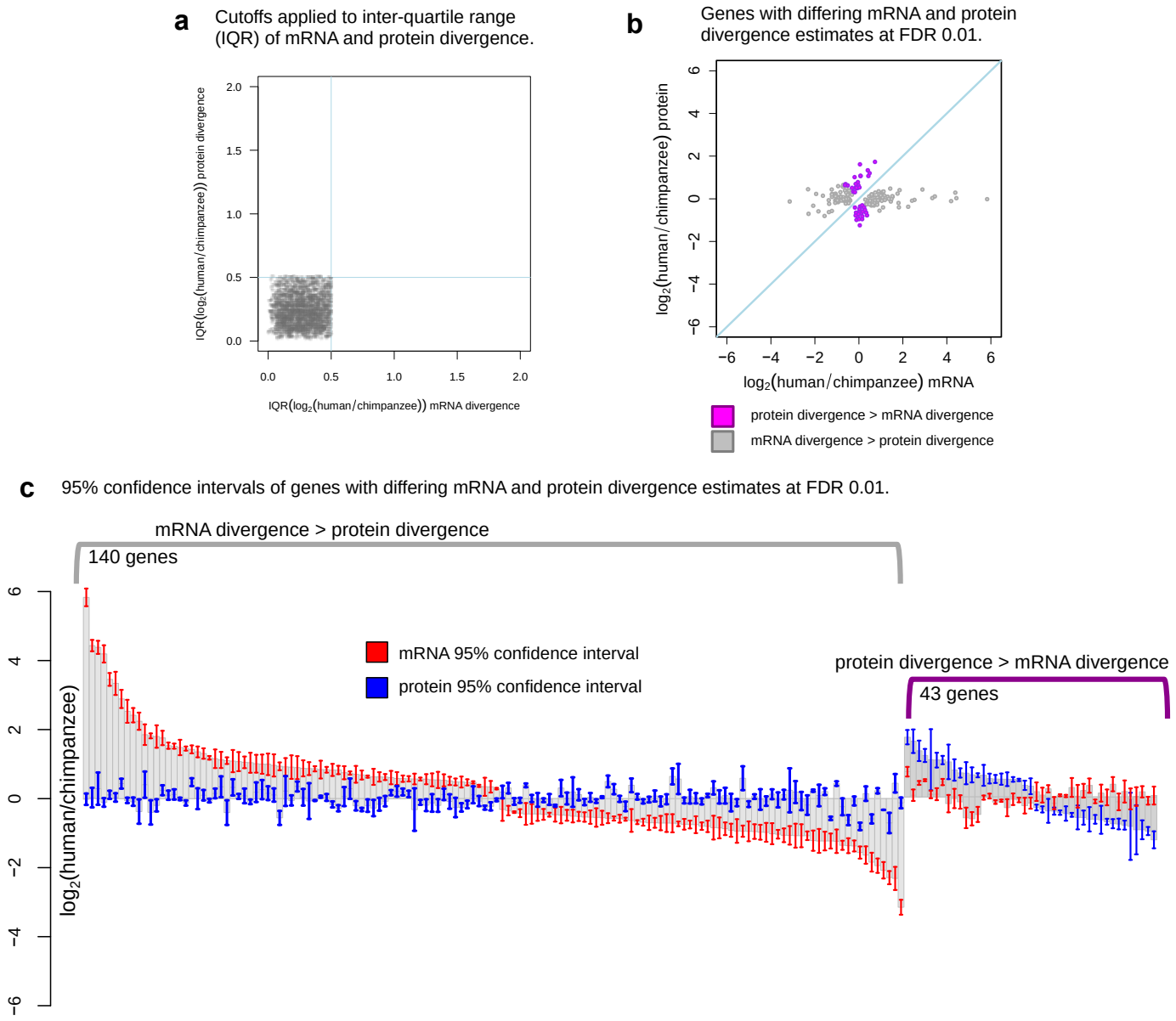


**Figure S9. Analysis of the subset of genes where the inter-quartile range (IQR) of mRNA and protein divergence is less than 0.7 and genes were quantified in all 5 individuals.** One possible explanation for our finding that protein expression levels are more conserved than mRNA expression levels could be more noise in protein measurements relative to mRNA measurements. To address this possibility, we applied a stringent filter to the genes we analyzed. We considered only genes where estimates of mRNA and protein divergene had an IQR of < 0.7. In (a) we provide a scatter plot of the IQRs of the genes we analyzed. Furthermore, we additionally required quantification in all 5 individuals in human and chimpanzee to assure IQRs were well estimated. This stringent filtering resulted in a subset of 1,981 genes (out of 3,390 genes) and eleminated cases were inconsistent measurements were obtained across human or chimpanzee individuals. (b) and (c) illustrate that our findings still hold true. Consistent with Fig. 1C and Fig. 1D, among genes with differing mRNA and protein divergence estimates, we observe a larger number of genes with higher absolute mRNA divergence than absolute protein divergence.

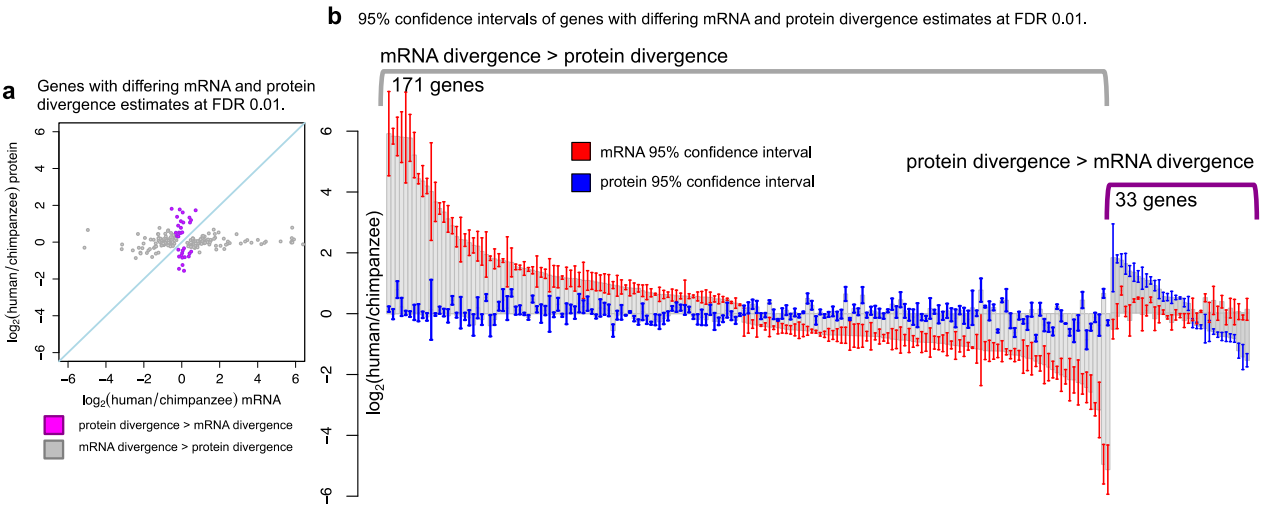


**Figure S10. Analysis of human and chimpanzee divergence estimates at the mRNA and protein level for genes quantified in all 5 individuals from both species.** One possible explanation of the relative stability of protein levels compared with mRNA levels is insufficient quantification of protein levels across individuals. The cases where protein levels are stable might be enriched for proteins quantified at our cutoff of 3 individuals. To address this possibility, we performed an alternative analysis whereby we required that proteins be quantified in all 5 individuals in both human and chimpanzee to eliminate proteins that might reflect insufficient quantification. This filtering left 2,651 genes (out of 3,390). We find that the results in (a) and (b) are consistent with Fig. 1C and Fig 1D. The number of genes where mRNA divergence exceeds protein divergence is higher than cases where protein divergence is larger than mRNA divergence.

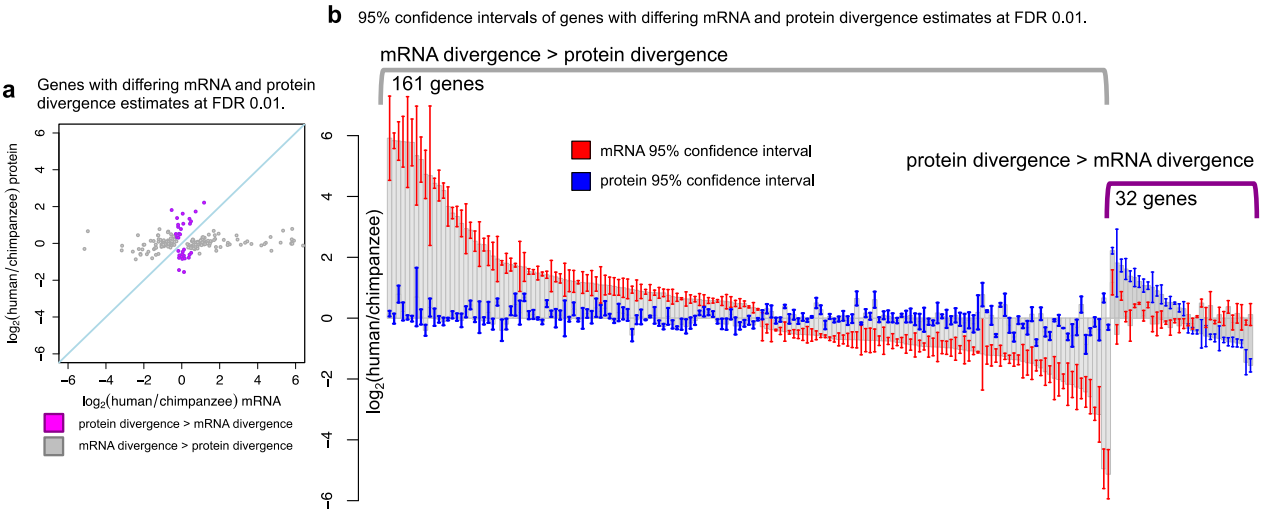


**Figure S11. Analysis of the subset of genes where proteins were quantified by at least two distinct peptides per individual.** Applications of SILAC often involve one or two biological replicates. Thus, at least 2 peptides are typically required per protein. Since our study required biological replication across species and individuals (proteins must be quantified in all species in 3 or more individuals), we allowed proteins to be quantified using a single peptide. We asked if this decision had any marked impact on our observations. To do so, we analyzed genes where proteins were quantified by at least 2 peptides in 3 or more individuals in all species. In total, there were 2,832 genes (out of 3,390). As illustrated in (a) and (b) above our results are robust with respect to this filtering. Consistent with Fig. 1C and Fig. 1D, among genes whose inter-species mRNA and protein divergence differ, we observe more genes with higher absolute mRNA divergence than genes with higher absolute protein divergence.

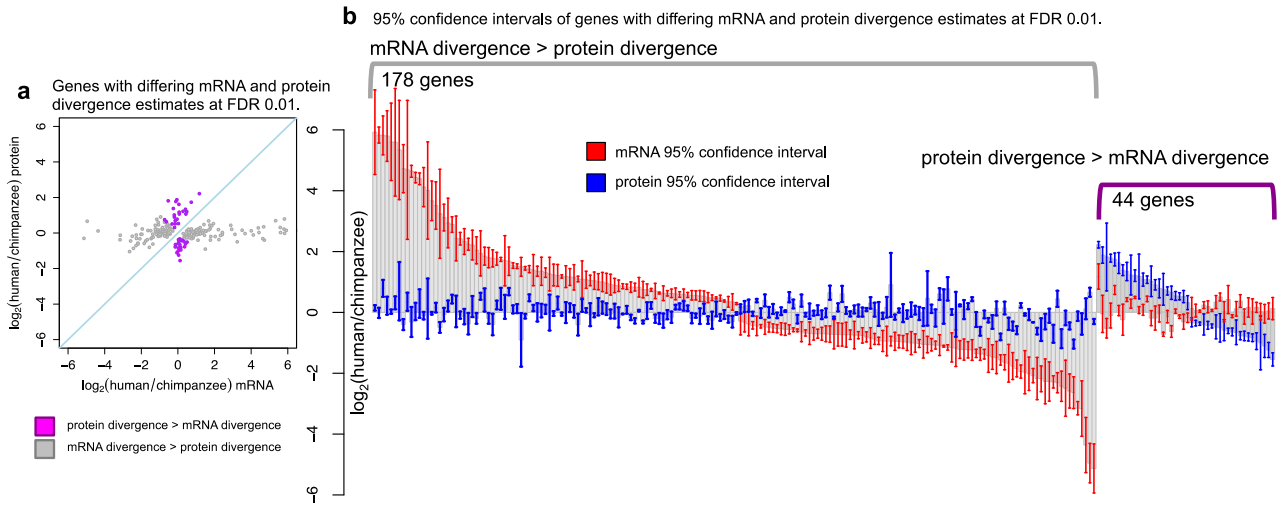


**Figure S12. Analysis of genes excluding the top 2% of most highly expressed genes in RPKM transcript levels.** If there are a few highly expressed genes that are differentially expressed (DE) between chimpanzee and human, then many genes expressed at intermediate or low levels might also appear as DE due to the finite sequencing depth and the dependent relationship between quantification levels using RNA-Seq. To address this possibility, we excluded the top 2% of highly expressed genes from each sample before we computed RPKM and normalized the $\log_2$(sample/standard) RPKM ratios from each individual relative to the internal standard LCL. If our results were due to finite sequencing depth, removal of these genes would greatly reduce the number of DE transcripts in our re-analysis. In contrast, we observe that among genes whose inter-species mRNA and protein divergence differ there are more genes with higher mRNA divergence than genes with higher protein divergence (consistent with the results in Fig. 1C and Fig. 1D).

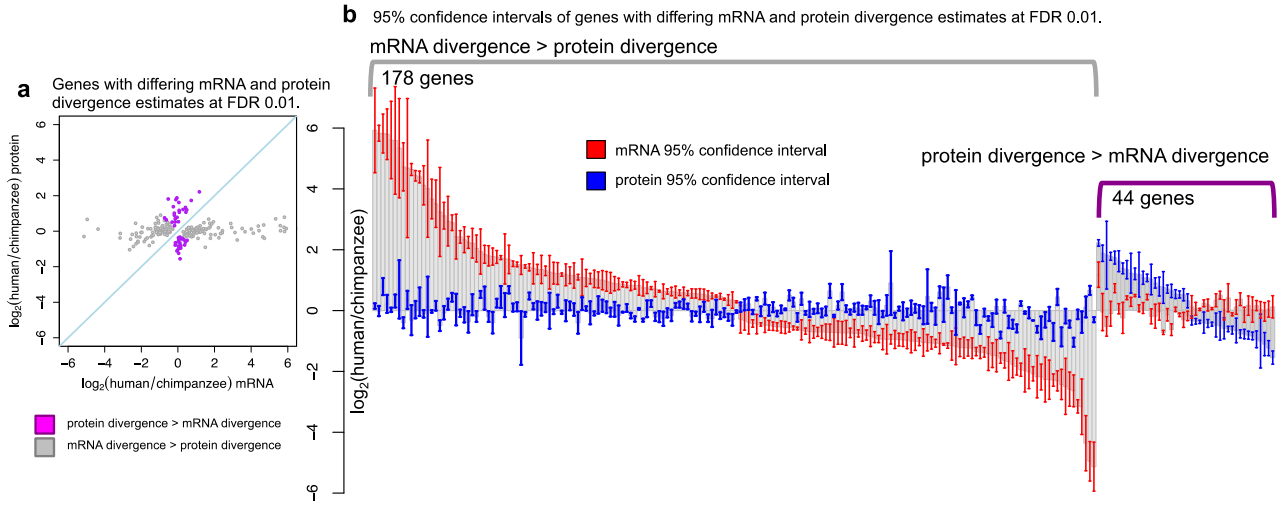


**Figure S13. Analysis of the subset of genes with RPKM > 1.** Previous studies have shown that RNA expression measurements of RPKM < 1 might often be unreliable (*28*). To rule out the possibility genes with small RPKM values might influence our observations, we removed any genes if the RPKM was less than 1 in any of the chimpanzee or human individuals. This filtering left 3,211 genes (out of 3,390). (a) and (b) illustrate that our results are robust with respect to this filtering. Consistent Fig. 1C and Fig. 1D, among genes whose inter-species mRNA and protein divergence differ significantly, we observe more genes with higher mRNA divergence than genes with higher protein divergence.

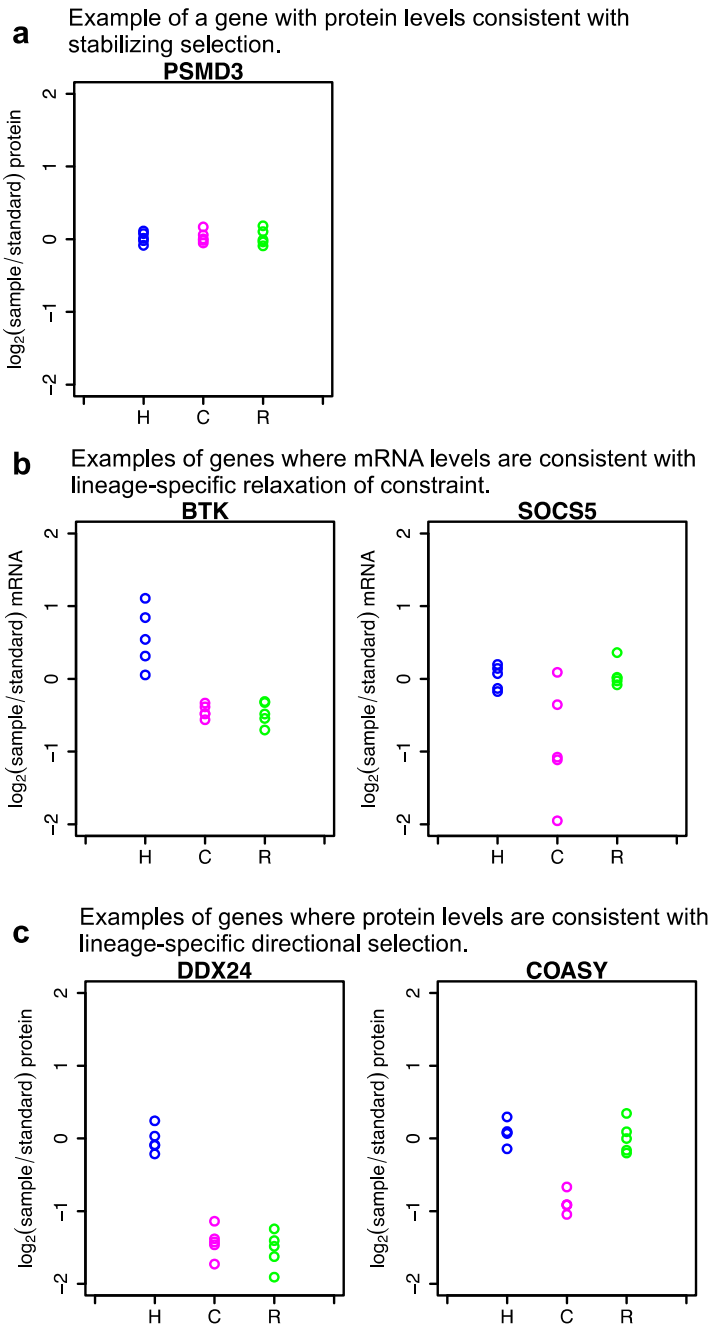


**Figure S14. Examples of protein and mRNA expression level patterns consistent with differing evolutionary scenarios.** Both protein and mRNA levels are normalized relative to the corresponding protein and mRNA expression levels in the internal standard cell line. (a) PSMD3 is a gene consistent with stabilizing selection. (b) Left plot shows a gene where mRNA levels consistent with lineage-specific relaxation of constraint along the human lineage. The right plot shows this same pattern along the chimpanzee lineage. (c) Examples of genes with protein levels consistent with lineage-specific directional selection along the human (left) and chimpanze (right) lineages.

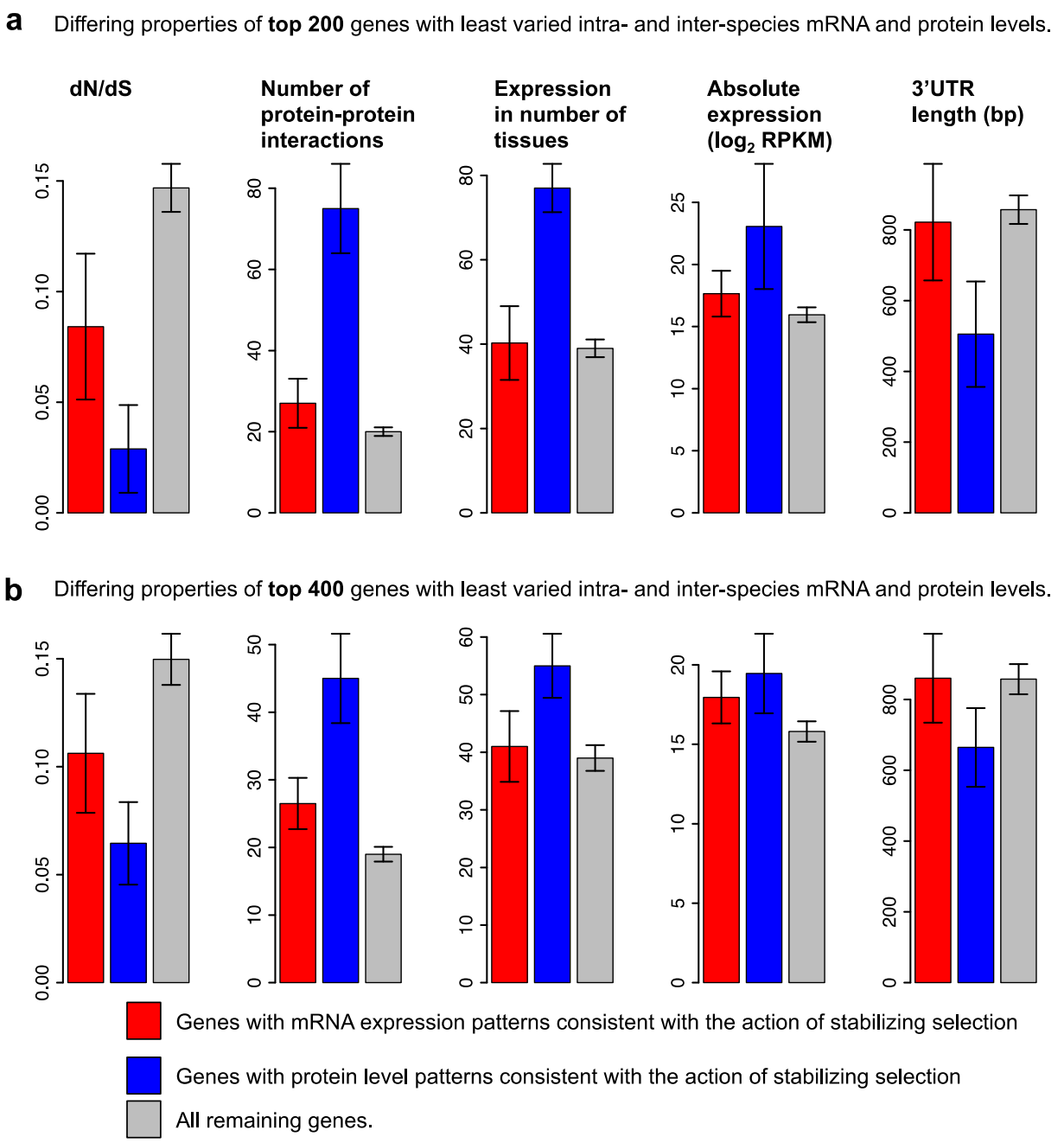


**Figure S15. Regulatory and sequence properties of genes with mRNA and protein levels consistent with stabilizing selection are robust with respect to the variance cutoff applied to this gene set.** (a) Properties computed using a strict cutoff applied to select genes consistent with stabilizing selection. (b) The same properties computed using a relaxed cutoff on the gene set. Intermediate values between these strict and relaxed cutoffs produced identical results. Error bars represent the 95% confidence intervals around the median**.**

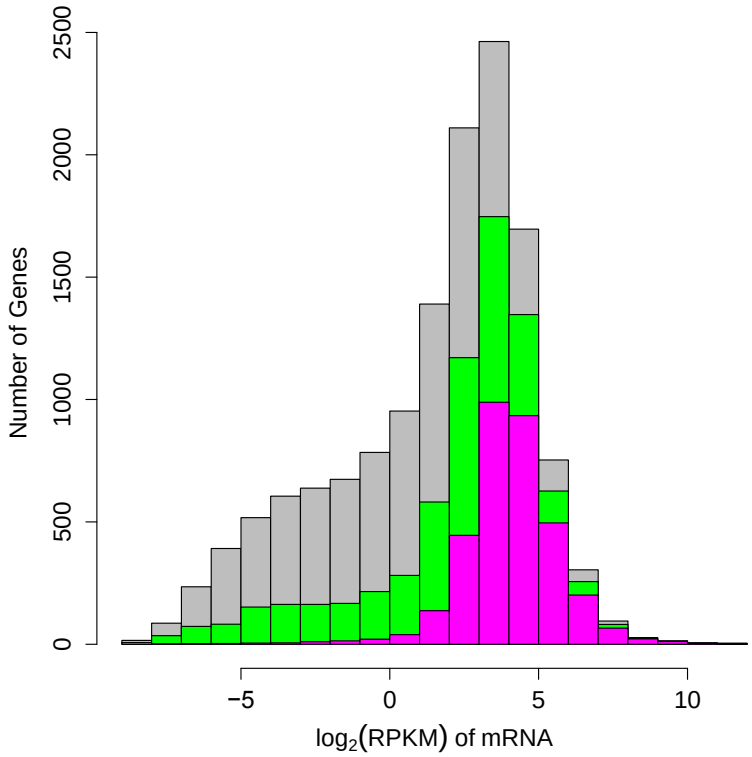


**Figure S16. RPKM values for transcripts quantified at the protein level.** Gray bars illustrate the RPKM distribution of all transcripts. Green bars designate the RPKM distribution of proteins detected and quantified in any one of the 15 cell lines. Magenta bars designate the RPKM distribution of proteins quantified in 3 or more individuals across all 3 species.

| Sample | Species | SDS-PAGE minigel fractions | LC Gradient Length | MS1 Spectra Collected | MS2 Spectra Collected |
|---|---|---|---|---|---|
| **GM18505** | Human | 40 | 1 hour | 62,497 | 473,147 |
| **GM18507** | Human | 40 | 1 hour | 91,999 | 422,500 |
| **GM18516** | Human | 40 | 1 hour | 110,208 | 405,340 |
| **GM19193** | Human | 40 | 1 hour | 72,124 | 382,871 |
| **GM19204** | Human | 40 | 1 hour | 70,686 | 447,980 |
| **18358** | Chimp | 40 | 1 hour | 79,024 | 388,336 |
| **18359** | Chimp | 40 | 1 hour | 60,238 | 486,655 |
| **3659** | Chimp | 40 | 1 hour | 80,787 | 401,536 |
| **4973** | Chimp | 40 | 1 hour | 88,420 | 355,404 |
| **Pt91** | Chimp | 40 | 1 hour | 75,775 | 406,529 |
| **150-99** | Rhesus | 40 | 1 hour | 87,195 | 377,597 |
| **R181-96** | Rhesus | 40 | 1 hour | 76,746 | 385,617 |
| **R249-97** | Rhesus | 40 | 1 hour | 94,467 | 402,710 |
| **265-95** | Rhesus | 40 | 1 hour | 63,536 | 465,632 |
| **R290-96** | Rhesus | 40 | 1 hour | 95,668 | 352,449 |

**Table S1. Raw Quantitative Mass Spectrometry Data.** A total of 1,209,370 MS1 spectra were collected and 6,154,303 MS2 spectra were collected. We used 600 gel fractions and 620 hours of instrument time. 143GB is the total data set size.

| Sample | Species | Quantification Events | Distinct Tryptic Peptides Quantified | Proteins Quantified |
|---|---|---|---|---|
| **GM18505** | Human | 94,156 | 52,809 | 5,085 |
| **GM18507** | Human | 83,664 | 49,308 | 4,853 |
| **GM18516** | Human | 76,105 | 43,573 | 4,787 |
| **GM19193** | Human | 72,293 | 45,408 | 4,633 |
| **GM19204** | Human | 87,193 | 50,664 | 4,971 |
| **18358** | Chimp | 68,565 | 43,049 | 4,531 |
| **18359** | Chimp | 98,653 | 57,959 | 5,142 |
| **3659** | Chimp | 75,682 | 45,376 | 4,765 |
| **4973** | Chimp | 59,969 | 37,133 | 4,283 |
| **Pt91** | Chimp | 75,943 | 45,944 | 4,755 |
| **150-99** | Rhesus | 58,011 | 33,150 | 3,915 |
| **R181-96** | Rhesus | 69,663 | 41,220 | 4,575 |
| **R249-97** | Rhesus | 62,602 | 35,547 | 4,218 |
| **265-95** | Rhesus | 81,883 | 46,247 | 4,552 |
| **R290-96** | Rhesus | 51,289 | 31,281 | 4,138 |

**Table S2. Quantified peptide and protein levels.** Peptide spectrum matches assigned at a false discovery rate of 1%. See methods for low-level peptide and protein quantification parameters.

| Sample | Species | Total Reads | High Quality Mapped | Excluded not exonic | Excluded Ambig. Mapping | Mapped to unique orthologous exons | % highQ mapped to genome/ total | % mapped to orthoexons/ highQ |
|---|---|---|---|---|---|---|---|---|
| **GM18505** | Human | 35,953,911 | 25,247,709 | 13,897,821 | 260,620 | 11,089,268 | 70.2% | 43.9% |
| **GM18507** | Human | 44,449,796 | 31,399,396 | 16,374,838 | 333,699 | 14,703,444 | 70.6% | 46.8% |
| **GM18516** | Human | 45,618,175 | 32,905,504 | 17,865,129 | 338,101 | 14,702,274 | 72.1% | 44.7% |
| **GM19193** | Human | 49,700,358 | 33,579,941 | 20,392,680 | 311,416 | 12,875,845 | 67.5% | 38.3% |
| **GM19204** | Human | 41,675,157 | 28,189,144 | 16,374,838 | 267,592 | 11,546,714 | 64.6% | 41.0% |
| **18358** | Chimp | 47,037,287 | 31,004,460 | 15628,483 | 338,903 | 15,037,074 | 65.9% | 48.5% |
| **18359** | Chimp | 45,626,484 | 30,097,227 | 15,114,805 | 343,478 | 14,638,944 | 66.0% | 48.6% |
| **3659** | Chimp | 38,911,057 | 26,393,759 | 13,509,525 | 290,421 | 12,593,813 | 67.8% | 47.7% |
| **4973** | Chimp | 44,699,355 | 29,829,994 | 15,168,988 | 316,341 | 14,344,665 | 66.7% | 48.1% |
| **Pt91** | Chimp | 41,641,279 | 26,549,913 | 13,872,803 | 277,749 | 12,399,361 | 63.8% | 46.7% |
| **150-99** | Rhesus | 41,032,986 | 25,884,588 | 13,350883 | 251,880 | 12,281,825 | 63.1% | 47.4% |
| **R181-96** | Rhesus | 34,437,003 | 22,026,194 | 10,180,184 | 231,121 | 11,614,889 | 64.0% | 52.7% |
| **R249-97** | Rhesus | 28,836,918 | 18,624,238 | 8,776,106 | 198,896 | 9,649,236 | 64.6% | 51.8% |
| **265-95** | Rhesus | 43,635,552 | 27,490,983 | 14,110,454 | 260,187 | 13,120,342 | 63.0% | 47.7% |
| **R290-96** | Rhesus | 37,046,951 | 24,005,747 | 11,119,048 | 251,425 | 12,635,274 | 64.8% | 52.6% |
| **GM19238** | Human internal standard | 40,211,719 | 28,637,456 | 15,711,461 | 277,891 | 12,648,104 | 71.2% | 44.2% |

**Table S3. Summary of the RNA-seq data.** 35GB of GZIP compressed RNA-seq data, 50bp unpaired reads, BWA quality score >10 (highQ above) from an Illumina Hi-Seq 2000. We excluded reads that did not map to orthologous exons or mapped ambiguously to more than one orthologous exon.

## The column descriptions in the Table S4 spreadsheet are below:

**ENSG**: Ensemble gene identifier.

**human.GM18505.protein, human.GM18507.protein, human.GM18516.protein, human.GM19193.protein, human.GM19204.protein, chimp.18358.protein, chimp.18359.protein, chimp.3659.protein, chimp.4973.protein, chimp.Pt91.protein, rhesus.150.99.protein, rhesus.R181.96.protein, rhesus.R249.97.protein, rhesus.265.95.protein, rhesus.R290.96.protein**: $\log_2$(sample/standard) protein ratios for human, chimpanzee, and rhesus samples, NA designates no protein level was quantified

**gene.symbol**: Gene symbol obtained from the UCSC human genome build hg19.

**N.H.protein, N.C.protein, N.R.protein**: number of protein quantifications obtained from the human(H), chimpanzee (C), and rhesus (R) samples. Note that genes quantified in 3 or more individuals in all 3 species as both mRNA and protein are included in this table.

**HC.pvalues.protein, HC.qvalues.protein, HR.pvalues.protein, HR.qvalues.protein, CR.pvalues.protein, CR.qvalues.protein**: P-value and corresponding FDR corrected Q-value for human-chimp (HC), human-rhesus (HR), and chimp-rhesus comparisons using protein levels.

**mean.H.protein, mean.C.protein, mean.R.protein**: mean $\log_2$(sample/standard) protein ratios computed from human (H), chimpanzee (C), and rhesus (R) individuals respectively.

**S.pvalues.protein, S.qvalues.protein**: P-value and corresponding FDR corrected Q-value for test of difference among the mean protein levels

**model.num.protein**: 1 = protein expression level pattern consistent with directional selection along human lineage, 2 = protein expression level pattern consistent with directional selection along chimpanzee lineage, 3 = undetermined pattern, 4 = patterns with no significant difference between mean expression levels; 5 = evidence for relaxation of constraint along human lineage, 6 = evidence of relaxation of constraint along chimpanzee lineage

**S.var.protein:** overall variance of protein expression levels
**S.var.rank.protein:** Rank based on overall variance among patterns with no significant difference between mean protein levels (model.num.protein=4). Note, genes that we consider consistent with stabilizing selection on protein levels, for purposes of our analyses, are those with model.num.protein = 4 and S.var.rank.protein $\leq$ 300.

**human.GM18505.rna, human.GM18507.rna, human.GM18516.rna, human.GM19193.rna, human.GM19204.rna, chimp.18358.rna, chimp.18359.rna, chimp.3659.rna, chimp.4973.rna, chimp.Pt91.rna, rhesus.150.99.rna, rhesus.R181.96.rna, rhesus.R249.97.rna, rhesus.265.95.rna, rhesus.R290.96.rna**: $\log_2$(sample/standard) mRNA ratios for human, chimpanzee, and rhesus samples, derived from RPKM ratios across orthologous exons, NA designates transcript could not be quantified in that sample

**standard.GM19238.rna.RPKM**: RPKM values computed from RNA-seq data derived from the internal standard sample from human LCL GM19238

**N.H.rna, N.C.rna, N.R.rna**: number of mRNA quantifications obtained from the human(H), chimpanzee (C), and rhesus (R) samples. Note that genes quantified in 3 or more individuals in all 3 species as both mRNA and protein are included in this table.

**HC.pvalues.rna, HC.qvalues.rna, HR.pvalues.rna, HR.qvalues.rna**, **CR.pvalues.rna, CR.qvalues.rna**: P-value and corresponding FDR corrected Q-value for human-chimp (HC), human-rhesus (HR), and chimp-rhesus comparisons using mRNA levels.

**mean.H.rna, mean.C.rna, mean.R.rna**: mean $\log_2$(sample/standard) mRNA levels computed from human (H), chimpanzee (C), and rhesus (R) individuals respectively.

**S.pvalues.rna, S.qvalues.rna: P**-value and corresponding FDR corrected Q-value for test of difference among the mean $\log_2$(sample/standard) mRNA levels.

**model.num.rna**: : 1 = mRNA expression level pattern consistent with directional selection along human lineage, 2 = mRNA expression level pattern consistent with directional selection along chimpanzee lineage, 3 = undetermined pattern, 4 = patterns with no significant difference between mean expression levels; 5 = evidence for relaxation of constraint along human lineage, 6 = evidence of relaxation of constraint along chimpanzee lineage

**S.var.rna:** overall variance of mRNA expression levels
**S.var.rank.rna:** Rank based on overall variance among patterns with no significant difference between mean mRNA expression levels (model.num.rna=4). Note, genes that we consider consistent with stabilizing selection on mRNA levels, for purposes of our analyses, are those with model.num.rna = 4 and S.var.rank.rna $\leq$ 300.

| Comparison | DE mRNA only | DE protein and DE mRNA | DE protein only |
|---|---|---|---|
| **Human-chimpanzee** | 580 | 235 | 336 |
| **Human-rhesus** | 876 | 781 | 668 |
| **Chimpanzee-rhesus** | 832 | 703 | 566 |

**Table S5. Genes classified based on differential expression (DE) at FDR 1%.**

| Term | Count | % | p-value | Fold Enrichment |
|---|---|---|---|---|
| GO:0030529~ribonucleoprotein complex | 120 | 40.7 | 1.56E-46 | 3.7 |
| GO:0006414~translational elongation | 58 | 19.7 | 3.30E-43 | 7.9 |
| GO:0022626~cytosolic ribosome | 51 | 17.3 | 2.46E-39 | 8.1 |
| GO:0003723~RNA binding | 106 | 35.9 | 2.86E-30 | 3.0 |
| GO:0003735~structural constituent of ribosome | 56 | 19.0 | 6.60E-29 | 5.4 |
| GO:0044445~cytosolic part | 52 | 17.6 | 4.26E-28 | 5.6 |
| GO:0033279~ribosomal subunit | 52 | 17.6 | 4.26E-28 | 5.6 |
| GO:0022625~cytosolic large ribosomal subunit | 29 | 9.8 | 2.77E-24 | 9.1 |
| GO:0006412~translation | 74 | 25.1 | 4.60E-24 | 3.4 |
| GO:0006396~RNA processing | 87 | 29.5 | 4.63E-23 | 2.9 |
| GO:0005840~ribosome | 58 | 19.7 | 7.18E-23 | 4.1 |
| GO:0005198~structural molecule activity | 59 | 20.0 | 4.83E-20 | 3.6 |
| GO:0015934~large ribosomal subunit | 30 | 10.2 | 6.65E-18 | 6.3 |
| GO:0008380~RNA splicing | 57 | 19.3 | 1.58E-16 | 3.2 |
| GO:0022627~cytosolic small ribosomal subunit | 22 | 7.5 | 2.03E-15 | 7.6 |
| GO:0006397~mRNA processing | 57 | 19.3 | 9.26E-15 | 2.9 |
| GO:0016071~mRNA metabolic process | 60 | 20.3 | 1.51E-14 | 2.8 |
| GO:0005681~spliceosome | 37 | 12.5 | 5.16E-14 | 4.0 |
| GO:0000398~nuclear mRNA splicing, via spliceosome | 40 | 13.6 | 1.22E-12 | 3.4 |
| GO:0000377~RNA splicing, via transesterification reactions with bulged adenosine as nucleophile | 40 | 13.6 | 1.22E-12 | 3.4 |
| GO:0000375~RNA splicing, via transesterification reactions | 40 | 13.6 | 1.22E-12 | 3.4 |
| GO:0031981~nuclear lumen | 101 | 34.2 | 5.03E-10 | 1.7 |
| GO:0005829~cytosol | 101 | 34.2 | 5.55E-10 | 1.7 |
| GO:0015935~small ribosomal subunit | 22 | 7.5 | 9.22E-10 | 4.6 |
| GO:0034470~ncRNA processing | 28 | 9.5 | 6.73E-08 | 3.1 |
| GO:0043232~intracellular non-membrane-bounded organelle | 125 | 42.4 | 7.33E-08 | 1.5 |
| GO:0043228~non-membrane-bounded organelle | 125 | 42.4 | 7.33E-08 | 1.5 |
| GO:0022613~ribonucleoprotein complex biogenesis | 32 | 10.8 | 9.33E-08 | 2.8 |
| GO:0042254~ribosome biogenesis | 25 | 8.5 | 1.85E-07 | 3.2 |
| GO:0006364~rRNA processing | 21 | 7.1 | 5.88E-07 | 3.5 |
| GO:0016072~rRNA metabolic process | 21 | 7.1 | 7.71E-07 | 3.5 |
| GO:0030532~small nuclear ribonucleoprotein complex | 11 | 3.7 | 1.81E-06 | 6.2 |
| GO:0034660~ncRNA metabolic process | 30 | 10.2 | 3.74E-06 | 2.5 |
| GO:0005654~nucleoplasm | 60 | 20.3 | 5.97E-06 | 1.7 |
| GO:0005730~nucleolus | 53 | 18.0 | 3.71E-05 | 1.7 |
| GO:0006413~translational initiation | 12 | 4.1 | 8.53E-05 | 4.0 |
| GO:0003743~translation initiation factor activity | 13 | 4.4 | 9.61E-05 | 3.7 |

**Table S6. Gene ontology enrichment analysis of the top 300, ranked by variance, protein expression level patterns consistent with the action of stabilizing selection at $P < 0.0001$.**

| Term | Count | % | p-value | Fold Enrichment |
|---|---|---|---|---|
| **GO:0007029~endoplasmic reticulum organization** | 4 | 2.1 | 0.001248203 | 15.4 |
| **GO:0015031~protein transport** | 31 | 16.1 | 0.003509713 | 1.7 |
| **GO:0045184~establishment of protein localization** | 31 | 16.1 | 0.003663969 | 1.7 |
| **GO:0008104~protein localization** | 32 | 16.6 | 0.004939603 | 1.6 |

| Term | Count | % | p-value | Fold Enrichment |
|---|---|---|---|---|
| **hsa00051:Fructose and mannose metabolism** | 7 | 4.5 | 4.40E-05 | 9.4 |
| **hsa00240:Pyrimidine metabolism** | 9 | 5.8 | 0.001394681 | 3.9 |
| **GO:0044275~cellular carbohydrate catabolic process** | 8 | 5.1 | 0.002160848 | 4.2 |
| **hsa00230:Purine metabolism** | 9 | 5.8 | 0.003365612 | 3.4 |
| **GO:0016779~nucleotidyltransferase activity** | 9 | 5.8 | 0.003697477 | 3.4 |
| **GO:0016052~carbohydrate catabolic process** | 8 | 5.1 | 0.004720748 | 3.7 |

**Table S7. Gene ontology and pathway enrichment analysis for protein levels consistent with the action of directional selection along the human lineage (top) and the chimpanzee lineage (bottom) at P < 0.005.**

**References and Notes**


**21.** M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 389, 1017 (Oct, 2007).

**22.** Z. Khan, J. S. Bloom, B. A. Garcia, M. Singh, L. Kruglyak, Protein quantification across hundreds of experimental conditions. *Proceedings of the National Academy of Sciences* 106, 15544 (September 15, 2009).

**23.** J. D. Storey, A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64, 479 (2002).

**24.** D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* 4, 44 (2008).

**25.** A. I. Su *et al.*, A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6062 (2004).

**26.** P. V. Hornbeck *et al.*, PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Research* 40, D261 (2012).

**27.** B. Schwanhausser *et al.*, Global quantification of mammalian gene expression control. *Nature* 473, 337 (2011).

**28.** D. Hebenstreit *et al.*, RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Molecular systems biology* 7, (2011).