Supplementary Materials

Figures S1 - S7
Tables S1 - S7
Extended Experimental Procedures
Supplementary References

Supplementary Figures

Figure S1. H3K27ac ChIP-Seq in human, rhesus and mouse embryonic limb, related to Figure 1. A. Human, rhesus, and mouse limb developmental stages interrogated in this study. Representative forelimb limb buds (E33 and E41 in human) and autopods (E44 and E47 in human) are shown for each indicated time point. **B.** *Upper panel:* Union of all H3K27ac regions identified in rhesus or mouse limb at one or more time points. *Lower panel*: Annotation at each time point. A schematic of limb morphology is shown for each stage (not to scale). A total of 68376 peaks were identified in rhesus. This number is the union of peaks called at E36 and peaks that reproduced between at least two early time points. All rhesus regions were converted to hg19 coordinates (66247 lifted coordinates) for annotation based on Ensembl v67. Numbers listed for each rhesus time point are from individual specimens, hence the larger number of peaks. The total number of peaks called on rheMac2 are listed below each chart, and the number of peaks lifted to hg19 are indicated in parentheses. Mouse peaks were annotated using Ensembl v67 natively on mm9 coordinates. Mouse values at each time point are based on reproducible peaks between two biological replicates. **C.** Enhancers active in human limb are enriched near genes with limb-specific expression. Limb specific enhancers were identified by k-means clustering of human limb E44 samples versus 7 human ENCODE cell-lines as in Figure 1B. Comparisons of global expression profiles of human limb E44 samples and the same ENCODE cell-lines as above identified 351 genes (Ensembl v67) with strong limb tissue specificity scores (TSPS) (Extended Experimental Procedures). Frequencies of limb-specific enhancers associated with limb-specific genes were plotted as red bars. Mean frequencies (from 1,000 simulations) of limb-specific enhancers associated with 351 randomly chosen genes were plotted in black bars, with the error bars representing 99.9% quantile values. **D.** H3K27ac identifies tissue-specific enhancers throughout mouse limb development. *Left Panel:* K-means clustering of H3K27ac signals across 153,845 putative enhancer regions identified in mouse limb or indicated mouse cell lines or tissue (k=16). The cluster of enhancers showing strong H3K27ac specifically in limb ("limb-specific" enhancers) is indicated by the bracket. *Right Panel:* Gene ontology and mouse phenotype enrichments calculated by GREAT (Mclean et al., 2010) for the strong limb-specific enhancer cluster. **E.** H3K27ac identifies temporal patterns of enhancer activation in the limb. *Left Panel.* K-means clustering identifies putative enhancers that show evidence of time point specific activation. All putative enhancers active in human limb were used in the clustering (25,436 enhancers, k=12), but only four clusters containing 6,768 enhancers that demonstrate strong time point specific H3K27ac signals are shown. *Right Panel.* Gene ontology and mouse phenotype enrichments calculated by GREAT (Mclean et al., 2010) for enhancers showing strong H3K27ac signal specifically in E33 or E47 limb (indicated on the heatmap by brackets).

Figure S2. Dynamic activation of the *HoxD* gene cluster during limb development is conserved between human, rhesus and mouse, **related to Figure 3**. *Top panel*: H3K27ac signal obtained from representative limb samples from each human time point at the *HOXD* cluster (hg19 coordinates). *Middle panel*: H3K27ac signal from the indicated mouse time points at the *HoxD* cluster (mm9 coordinates). *Bottom panel*: H3K27ac signal from the indicated rhesus time points at the *HoxD* cluster (rm2 coordinates).

Figure S3. Multi-species Analysis of Limb H3K27ac ChIP-Seq, related to Figure 3. A. Data analysis pipeline. **B.** Principal component analysis of human and mouse H3K27ac signals in limb, embryonic stem cells and adipocytes (Mikkelsen et al., 2010; Shen et al. 2012).

Figure S4. Identification of human lineage gains of H3K27ac, related to Figure 4. A. Conserved H3K27ac profiles at two canonical limb genes, *PITX1* and *TBX5*. *Left panel*: H3K27ac signal obtained from representative limb samples from each human time point at *PITX1* (hg19 coordinates). Green horizontal bars indicate H3K27ac enriched regions. Lighter and darker green bars indicate enriched regions identified at the indicated rhesus or mouse time points. Coordinates for human, rhesus, and mouse are respectively based on the hg19, rheMac2, and mm9 assemblies. *Right panel*: H3K27ac signal from the indicated human, rhesus, and mouse time points at *TBX5*. **B.** Distributions of length changes of the rhesus and mouse orthologous sequences for all 3-way orthologous sequences (grey) and human gains (red). **C.** Significant keyword enrichments from annotation clusters reported by DAVID (Dennis et al., 2003) for human gain promoters. **D.** Fold enrichment determined by QPCR of human E44 limb ChIP-enriched material for the indicated amplicons relative to ChIP-enriched material from E11.5 and E13.5 mouse limb, normalized for input. Error bars indicate mean +/- SEM. Regions not enriched in E44 samples relative to input are indicated by *.

Figure S5. Evolutionary origins of sequences associated with human lineage gain of activation, related to Figure 5. A. Age and constraint analysis of promoters and enhancers active in human limb. *Left four panels*: Distribution of age estimates for all orthologous, stably marked, and human gain enhancers or promoters at each human time point, mapped onto the known phylogeny of vertebrate genomes used in the analysis. We estimated the age of each enhancer by identifying the most distant vertebrate lineage with an orthologous sequence (Extended Experimental Procedures). Estimated ages (in millions of years) for internal nodes in the phylogeny are shown. In the boxplots, the right, middle and left bars of the boxes represent the 25^{th}, 50^{th} and 75^{th} percentiles of the data, respectively. Whiskers extend from the box to the most extreme data point that is less than 1.5 times the interquartile range. The human gain enhancers are significantly younger than the 3-way orthologous and stably marked enhancers at all time points (Wilcoxon rank-sum $p<0.05$). The human gain promoters at E44 and E47 are significantly younger than the 3-way orthologous and stably marked promoters (Wilcoxon rank-sum $p<0.05$), while no significant difference is identified for promoters

at E33 and E41. (***, p<0.0001. **, p<0.001. *, p<0.05). *Right four panels*: Proportions of promoter and enhancer regions classified as above that are conserved at each indicated subtree node or the whole tree using phyloP likelihood ratio test (Extended Experimental Procedures). Significant differences in proportion of 3-way orthologous and stably marked regions compared to human gain regions are indicated by * (Fisher's exact p <0.01). **B.** Inferred human gain mechanisms based on H3K27ac marking. *Modification*: human gain orthologous sequences marked by H3K27ac in rhesus or mouse limb, using data generated in this study. *Co-option*: human gain orthologous sequences marked in at least one of 18 mouse non-limb tissues or cell-lines (Shen et al., 2012). *De novo*: human gain orthologous sequences not marked in any of the rhesus and mouse tissues/cell-lines investigated.

Figure S6. Human gain enhancers do not exhibit a significant human-specific increased substitution rate compared to other primates, related to Figure 5.

A. Upper box plots show distributions of human-specific substitution rates, chimpanzee-specific substitutions rates, ratio of human-specific to chimpanzee-specific substitutions, ape-specific substitution rates, rhesus-specific substitution rates and ratio of ape-specific to rhesus-specific substitutions for the indicated enhancer classes at human E44. The lower table contains results for difference between human gain versus stably marked regions at all time points, with the values in each cell representing mean difference (upper value) and Wilcoxon rank-sum P values (lower value). Ratios of human-specific to chimpanzee-specific substitutions, or ape-specific to rhesus-specific substitutions were not significantly different between human gain and stably marked regions. B. Same as A, for promoter analysis.

Figure S7. Potential functional outcomes of human lineage gain of regulatory activity, related to Figure 6. A. Human gain enhancers are significantly enriched near genes with ≥4-fold increases in expression level in human E44 limb compared to mouse E11.5 limb. Enhancer peaks in human were assigned to all human Ensembl genes as described (Mclean et al., 2010). Human-mouse 1:1 orthologs were binned by the magnitude of differential human expression versus mouse. Frequencies of human-gain enhancer peaks assigned to human genes are plotted as red bars. Mean frequencies (from 1000 simulations) of the same number of randomly sampled stably marked orthologous enhancers associated with genes in each bin were plotted as grey bars (error bars representing 0.1% and 99.9% quantile values). **B.** Limb H3K27ac signal from the indicated human, rhesus, and mouse time points at the promoter (left) and a potential enhancer (right) of the *ARGHAP6* gene. Horizontal bars indicate regions of H3K27ac enrichment. Coordinates for human, rhesus, and mouse are for hg19, rheMac2, and mm9 genomes respectively. Vertical green bars indicate total signal in human region and relative signal in orthologous regions from rhesus and mouse; P values indicating significant human lineage increased marking are shown. Vertical black and grey bars indicate relative *ARGHAP6* expression determined by E44 and E11.5 RNA-seq. **C.** Left panel: a normal adult hindlimb skeleton. Right panel: an *Xpl* mutant hindlimb skeleton. The locations of digits (numbers), fibula (Fi), and tibia (Ti) are indicated on each image. Abnormal digits in the *Xpl* mutant are indicated by *. Hindlimb digit 1 in the *Xpl* mutant

is noticeably thickened and elongated. Reprinted from Masuya H, Sagai T, Moriwaki K, Shiroishi T. 1997. Multigenic control of the localization of the zone of polarizing activity in limb morphogenesis in the mouse. Dev Biol 182: 42–51w ith permission from Elsevier.

Supplementary Tables

Table S1
: Sequencing statistics and H3K27ac enriched regions summaries, related to Figure 1.

Table S2
: H3K27ac read counts in enriched regions and significance of species differences for all comparisons performed, related to Figures 3 and 4.

Table S3
: H3K27ac enriched regions identified as human gains in bed format; qPCR primers for human and mouse orthologous regions used to confirm human gain of H3K27ac in E44 limb; sequences marked by H3K27ac in human limb that are species or lineage-specific; contingency table for human versus rhesus pairwise comparisons, related to Figures 4 and 5.

Table S4
: H3K27ac enriched regions identified as human losses in bed format, related to Figure 4.

Table S5
: Transcription factor binding motif and repeat content comparisons for human gain regions versus stably marked regions; contingency table for human versus rhesus or mouse motif occurance comparisons, related to Figures 4 and 5.

Table S6
: Orthologous genes between human and mouse associated with at least one human gain of H3K27ac, related to Figures 4 and 6.

Table S7
: Human gains that overlap human-accelerated non-coding sequences, related to Figure 7.

Extended Experimental Procedures

Tissue Collection, ChIP-seq and RNA-Seq

Human embryonic limb tissue was collected, staged and provided by the Joint MRC/Wellcome Trust Human Developmental Biology Resource (www.hdbr.org). Tissues were flash frozen upon collection and stored at -80 C. Human limbs were staged using the Carnegie staging system, which uses morphological landmarks including appearance of the limb buds, digital rays, and invagination of interdigital webbing (O'Rahilly et al., 1987). The use of human embryonic tissue in this study was reviewed and approved by the Yale Human Investigation Committee. For each time point replicate a single forelimb and single hindlimb autopod from the same embryo were combined, homogenized, and crosslinked as described for mouse tissue (Cotney et al., 2012). Fetal rhesus limb tissue from gestational days 31 to 36 was harvested according to approved Yale IACUC protocols (Dominguez et al., 2012). Single forelimb and single hindlimb rhesus autopods were combined as for human experiments. Mouse embryos were harvested in accordance with approved Yale IACUC protocols, and all mouse ChIP experiments were performed as previously described (Cotney et al., 2012). Only autopod tissue from forelimb and hindlimb were used for E12.5 and E13.5 experiments.

For all experiments, between 5 and 30 µg of total chromatin were incubated with 2 µg of H3K27ac antibody (Abcam ab4729) bound to protein G Dynabeads overnight with rotation at 4°C. Beads were washed and chromatin was eluted and purified as previously described[6]. Purified chromatin was prepared for sequencing using the standard Illumina multiplexing protocol. 75bp single end sequencing was performed for all samples on Illumina HiSeq 2000 instruments at the Yale Center for Genome Analysis (YCGA).

For RNA-Seq experiments, digit one was separated from remaining digits of E44 forelimb or hindlimb autopod tissue and placed directly in Qiazol. RNA was extracted from all tissue using miRNEasy kit (Qiagen) and quality was measured using an Agilent Bioanalyzer RNA pico chip (All RIN >8). RNA was prepared for sequencing using standard Illumina RNA-Seq library preparation protocols. 75bp paired end sequencing was performed on an Illumina GAIIx instrument at YCGA.

ChIP-seq read alignment and peak detection

ChIP-seq reads were aligned to respective genome references (hg19 for human, rheMac2 for rhesus, mm9 for mouse) using Bowtie (v0.12.7) (Langmead et al., 2009), keeping only uniquely mapped reads (-m1 option). A genome reference was built using fasta sequences of each species downloaded from UCSC Genome Browser (http://genome.ucsc.edu/ (Kent et al., 2002)), including all autosomes, sex chromosomes and mitochondrion sequences. Duplicated reads (same strand and same start site) were further filtered out for peak detection and signal generation. A sliding window approach was used to identify peaks in ChIP-seq data (Mikkelsen et al., 2010). A sliding window of 500 bp with a 25 bp step-size was used to determine the enrichment of aligned reads from each H3K27ac ChIP-seq sample. For each window, the total number of reads from H3K27ac and input control experiments were counted respectively. Raw read number in the input control experiment was then scaled to match the sequencing depth of each H3K27ac experiment. The number of expected reads in a window based on the

assumption of uniform distribution of total mapped reads along a specific chromosome was calculated. Significance of enrichment was calculated using a Poisson model where the null model is a larger number of input control read counts or expected read counts in the window. Significantly enriched windows (p-value$\leq 10^{-5}$) within 1kb of one another were then merged into a single region. Merged regions that did not overlap 1kb upstream of a transcription start site or exons in Ensembl (v67) annotation were denoted as candidate enhancers. ChIP-seq fragment densities were generated by extending each aligned read to 300 bp based on sonication size then counting the number of extended fragments that overlap each nucleotide. Read counts were then normalized to fragments per million aligned reads. Reproducible enriched regions were defined as those that had 1bp minimum overlap between two H3K27ac ChIP-seq replicates for each time point. Merged coordinates from both replicates were then used to define a reproducible region.

RNA-seq read alignment and expression quantification

The reference for aligning RNA-seq data was built using each genome reference (hg19 for human, mm9 for mouse) plus a custom splice junction library. The splice junction library, consisting of 140bp sequences that extended 70bp from the splice site to either side of annotated exons, was constructed using RSEQtools (Habegger et al., 2011). RNA-seq reads were aligned using Bowtie (v0.12.7) (Langmead et al., 2009), and only uniquely mapped reads were retained (-m1 option). RSEQtools was used to derive RPKM (Mortazavi et al., 2008) values for composite gene models built from Ensemblv67 annotations. A composite gene model is the union of bases covered by all transcripts annotated for a gene. For mouse RNA-seq data, the average RPKM values from forelimb and hindlimb were used to represent E11.5 expression. In order to compare human RNA-seq data to mouse, human digit one read counts were weighted 1/5 and digit 2 through digit 5 weighted 4/5 to derive expression values for each replicate. The average RPKM values from the processed data of two replicates were used to represent E44 expression.

Cross-species H3K27ac region comparison

H3K27ac regions were mapped between species using the liftOver tool and chain files from the UCSC Genome Browser. A region was considered as 2-way orthologous if all following criteria were satisfied: mapping from query to target species is unique, the reciprocal mapping from target species back to query is unique, and the coordinate derived after reciprocal mapping has at least 50% overlap with the original peak coordinate. Such 2-way orthologous regions were further filtered so that the coordinates in target species should be on autosomes, sex chromosomes and mitochondrion sequences (i.e. not on random chromosomes that were excluded in the alignment reference). 3-way orthologous regions were then selected as the intersection of 2-way orthologous regions between query species and two target species. (e.g., for human regions, human is the query species while rhesus and mouse are target species.) We restricted ourselves to human gain or human loss region detection to these 3-way orthologous regions because the orthology is clear and short-read mapping to these regions is comparable between species. We do not provide estimation on sequences that do not have clearly identifiable orthologs across species, for which our method is not suited. We also recognize that species or lineage-specific sequences could contribute to

novel enhancer function, and we provide an estimate of the frequency of these events (Table S3).

We detail below the quantitative strategy used for H3K27ac signal comparison.

Human gain regions

For reproducible regions detected at each time point in human, a subset of human gain regions were defined. For each human time-point, the signal was compared to all rhesus and mouse data in our time-series.

Pair-wise comparisons were made between a single human time point and all rhesus time points. To quantitatively compare the signal level of marking, we used a two-sided Fisher exact test for human and rhesus comparisons, as only one replicate was available for each rhesus time point. Reads from human time point replicates were pooled then used directly, while read counts in rhesus were scaled by a ratio of (human peak length)/(orthologous rhesus region length). A two-way contingency table was constructed for each region (Table S3).

Multi-sample comparisons were made between two replicates at a single human time point and two replicates at each mouse time point. Reads from human time point replicates were used directly, while read counts in mouse were scaled by a ratio of (human peak length)/(orthologous mouse region length). For these comparisons we adopted a published log linear model with a Poisson link and a subsequent likelihood ratio test based on model fit for comparison with replicates (Bullard et al., 2010; Cotney et al., 2012). The number of reads in a region p in sample i is denoted as X_{pi}, and is modeled as $\log\left(\mathrm{E}\left[X_{pi}|X_{.i}\right]\right) = \log X_{.i} + \lambda_{pj(i)} + \theta_{gi}$,where $X_{.i}$ is the total number of mapped reads for sample i, $\lambda_{pj(i)}$ is the species-specific signal level and θ_{gi} is the individual replicate effect. In the likelihood ratio test we evaluate which of two models better fit the data. For model one there exists a significant difference in signal between two species, while in model two there is no difference between two species. Resulting p-values were then adjusted using Benjamini-Hochberg method (bhP) (Benjamini and Hochberg, 1995).

Human gain regions were then defined as the intersection of significantly stronger human regions (bhP<=0.001 and signal fold change>=1.5) in all eight comparisons for each time point. From this analysis we also defined stably marked H3K27ac regions. These represent a subset of 3-way orthologous regions that exhibit similar levels of H3K27ac signal across all species. Specifically, these are regions that show no significant change in H3K27ac signal (signal fold change < 1.5 and > 2/3) between one time point in human and at least one time point in both rhesus and mouse.

Human loss regions

For reproducible regions detected at each time point in mouse experiments, a subset of human loss of H3K27ac were defined. We compared signal quantitatively using the same tests as above. However, since we can not match human and mouse time points perfectly, we require that human signal at a region orthologous to the region identified in mouse be significantly lower in all human time points to be considered human loss. Hence, 4 multi-sample comparisons between a single mouse time point and all human

time points and 16 pair-wise comparisons between all possible pairings of human and rhesus time points were conducted. Human loss regions were then defined as the intersection of human significantly weaker regions (bhP<=0.001 and signal fold change<=(2/3)) in all 20 comparisons.

Additional notes

Since matching developmental stages across species is challenging, we did not try to associate one human gestational stage with one certain stage in rhesus or mouse, and compare H3K27ac signals only for such "matched" time points. Instead, we compared each time point in human with all other time points in rhesus and mouse, so that we identify human lineage gain sites that display significantly less H3K27ac signal in rhesus and mouse across a range of comparable developmental stages. This approach is especially important to mitigate the danger of over-estimating human lineage gain regions due to lack of replicates for rhesus samples, by comparing each human time point to all four rhesus stages. We do note that the procedure for calling human specific gain or loss was conservative, and did not consider possible heterochrony in marking across species. However, these sets of peaks indicated a clean set of human absolute gain or loss for the time-series generated by this study, yielding high quality data for downstream analyses.

Validation of human specific H3K27ac regions

Human gain H3K27ac regions at E44 were chosen for testing that had an average RPKM ≥ 1 and when lifted to the mouse genome changed less than 10% in overall length to ensure comparable QPCR data. 300bp of human DNA sequence surrounding the peak of signal within each enriched region were selected and QPCR amplicons were designed using BatchPrimer3. To ensure similar QPCR efficiencies these 300bp regions were lifted to the mouse genome and QPCR amplicons were designed as above. 28 primer pairs and one positive and negative control for both species produced reliable QPCR products and were utilized for confirmation of human H3K27ac gain (Table S3). H3K27ac ChIP material from two independent experiments each of human E44, mouse E11.5, and mouse E13.5 limbs was tested for increased human H3K27ac. Human gain was determined by normalizing human E44 H3K27ac ChIP C_t values versus input C_t values in a fashion similar to □□C_t values typically utilized in detecting gene expression differences via RT-QPCR. P-values were calculated by 2-way ANOVA with Prism v5 (GraphPad).

Comparison to human and mouse ENCODE data

Raw reads from ChIP-seq and RNA-seq experiments in 7 ENCODE cell-lines (Gm12878, H1hesc, Helas3, Hepg2, Huvec, K562, Nhek) were downloaded from the UCSC Genome Browser. H3K27ac ChIP-seq and control data were generated at the Broad Institute (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeBroadHistone/) . RNA-seq data were generated at the California Institute of Technology (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeCaltechRnaSeq/). Single end 75bp data were then processed in the same fashion as above. Aligned reads

from ChIP-seq experiments in 18 mouse tissues/cell-lines (bone marrow, cerebellum, cortex, embryonic brain, adult heart, embryonic heart, adult liver, embryonic liver, intestine, kidney, lung, mMEF, mESC, olfactory bulb, placenta, spleen, testis, thymus reported by Shen et al. 2012) were downloaded from the UCSC Genome Browser. H3K27ac enriched regions were then identified in each sample as above.

K-means clustering of H3K27ac regions

A composite enhancer annotation for cross tissue comparisons was created by merging all enhancer regions identified in limb and ENCODE samples with a minimum of 1bp overlap criteria using BEDTools (Quinlan and Hall, 2010). To generate comparable signal values for each enhancer, reads from multiple replicates for each tissue or cell type were combined. Then total number of reads from each tissue or cell type was calculated for each enhancer region, such that each element was represented by a vector with a length of total tissues/cell-lines considered. Values for each element were normalized by subtracting the mean of tissue values from each individual tissue value and dividing by the standard deviation of values for that region. These normalized values were subjected to k-means clustering using R (http://cran.r-project.org/).

Enrichment of limb-specific enhancers near limb-specific genes in human

For each Ensembl gene, tissue specificity scores (TSPS) (Ravasi et al., 2010) were calculated using RPKM values in 8 tissue/cell-types. TSPS was computed as relative entropy: $TSPS_i = \sum_j f_j^i log_2 f_j^i / (q^i)$, where f_j^i is the RPKM of gene i in tissue/cell-type j divided by sum of RPKM values for gene i in all tissue/cell-types; q^i is the expected value of f_j^i with uniform expression assumption, and 1/8 specifically for this comparison. 351 limb-specific genes were selected based on the following criteria: the gene has its highest RPKM value in limb compared to human ENCODE data; it has an RPKM value greater than 10 in limb; it has a TSPS greater than 1.5. Merged H3K27ac regions in 8 tissue/cell-types were assigned to Ensembl genes using default association rules used by GREAT (basal plus extension; see great.stanford.edu). The gene regulatory domain is extended in both directions to the nearest gene's basal domain but no more than the maximum extension (1000kb) in one direction (Mclean et al., 2010). For the 351 limb-specific genes, limb-specific enhancers assigned to them were extracted and frequencies were counted from distance 1kb to 201kb at 10kb intervals. Then 1000 sets of 351 randomly selected Ensembl genes were generated, and the frequency of human gain enhancer association was counted as above for each set.

Comparison to human and mouse H3K27ac ChIP-seq data from ES and adipocytes

H3K27ac data from human and mouse adipocytes were obtained from Mikkelsen et al. (Mikkelsen et al., 2010), processed as described above. To generate the PCA plot in Figure S3B, regions identified in all human limb, ES and adipocyte data were first merged, and only 2-way orthologous regions were then used for signal comparison using R.

Conservation and human constraint

Per-base placental mammal PhastCons (Siepel et al., 2005) scores for the hg19 assembly were downloaded from the UCSC Genome Browser and used to assess conservation. Human population variation data from 1000 Genomes project phase 1 release (1000 Genomes Project Consortium et al., 2012) were used to evaluate human constraint. Only SNP variation data in low coverage samples were used and heterozygosity was calculated as 2*AF*(1-AF) (AF is allele frequency) (Mu et al., 2011). As the heterozygosity score increases the level of constraint in human decreases. For each element, bases were masked where no multiZ alignment was available, no DNA sequence variation could be called in 1000 Genomes project phase 1, or was indicated as gap in the hg19 build of the human genome. For each element, the mean unmasked base PhastCons score or human heterozygosity score for autosomal regions were calculated. Wilcoxon rank-sum test was used to assess the significance of a shift in per element mean phastCons or heterozygosity values among different classes of human enhancers established above. The background intergenic/intronic regions were extracted using hg19 sequence and Ensembl v67 annotation, defined as the masked genome minus exons and promoters. In figure 5A, per base average values were plotted for this background. To compare human gain enhancers and the intergenic/intronic background, 1000 random sampling iterations were conducted. For human gain enhancers in each time point, sequences matching the length distribution of such enhancers were randomly sampled from the background, which were then masked and calculated for per element mean phastCons or heterozygosity values as above. The values from random samples were then compared to those of human gain enhancers. Human gain enhancers show significantly reduced overall genetic diversity compared to background sequences in this analysis ($P < 0.007$), but do not show a significant difference compared to stably marked enhancers. However, as noted in the main text ascertainment biases and false positives in 1000 Genomes data may influence this result, which precludes us from drawing a strong conclusion regarding the level of constraint of human gain enhancers within human populations.

Cross-species H3K27ac marking and gene expression comparison

H3K27ac ChIP-seq and RNA-seq data in human E44 and mouse E11.5 were used to assess the association of gene expression changes and H3K37ac marking dynamics at one-to-one orthologous genes between human and mouse Ensembl v67 genes. A subset of orthologous genes were further selected for promoter region analysis: the human gene should have a H3K27ac peak that overlap sequence 1kb upstream of its TSS, the associated H3K27ac peak itself is 2-way orthologous between human and mouse, and the TSS of orthologous mouse composite gene model falls within the orthologous H3K27ac peak in mouse. The fold change of H3K27ac promoter peak marking strength was compared to the fold change of expression level in human versus mouse.

Enhancer peaks in human were assigned to all human Ensembl genes using GREAT default criteria (Mclean et al., 2010). All one-to-one orthologous genes were then binned by human versus mouse expression fold change by calculating the ratio of RPKM values for each one-to-one orthologous gene. For each bin, the number of human-gain enhancer peaks assigned to the human gene was counted. The same number of enhancers was then randomly sampled 1000 times from stably marked E44 enhancer regions and the frequency of associations with a human gene was counted for each set.

Motif analyses

532 transcription factor (TF) motifs were derived from the JASPAR database (Sandelin et al., 2004). Motif occurrences in each respective genomes were identified using FIMO (P value = 1e-5 or less); (Grant et al., 2011), Only 492 motifs had significant occurrences in the three species, and were used for further analysis. To investigate the possibility that a specific TF may contribute significantly to H3K27ac marking dynamics across species, the following procedure was adopted. For each motif, a two-way contingency table was constructed to assess the association (Table S5). The value in each cell was the number of H3K27ac regions that satisfied its corresponding row and column conditions. A two-sided Fisher exact test was used to determine the significance of association. In addition to individual motifs, combinations of two motifs were examined where the number of elements with any two specific motifs were counted in the contingency table. No significant enrichment was observed (data not shown).

Repeat content analyses

Repeat content annotations were downloaded from RepeatMasker tracks in the UCSC Genome Browser. To investigate the possibility that a specific repeat element may contribute to the origin of human specific gain enhancers, a two-sided Fisher exact test was performed on human specific gain and stably marked three way orthologous enhancers based on the same contingency tables constructed for motif analysis. Resulting p-values were then adjusted using Benjamini-Hochberg correction (Benjamini and Hochberg, 1995).

Substitution analyses

We considered two pairs of species or lineage-specific substitutions that may contribute to human gain H3K27ac regions: human-specific versus chimpanzee-specific substitutions and ape-specific versus rhesus-specific substitutions. Human-specific substitutions were defined as positions that differ in human from a conserved nucleotide in chimp, orangutan, rhesus, and marmoset. Chimpanzee-specific substitutions were defined as positions that differ in chimp from a conserved nucleotide in human, orangutan, rhesus, and marmoset. Ape-specific substitutions were defined as positions that differ in an ape consensus from a conserved nucleotide in rhesus, marmoset, and mouse. Rhesus-specific substitutions were defined as positions that differ in rhesus from a conserved nucleotide in human, chimp, orangutan, marmoset, and mouse.

The substitution rate for each element is defined as (# A-specific substitutions)/(# A-specific substitutions + # conserved nucleotides) where A is a substitution specific to human, chimp, ape or rhesus and the conserved nucleotide is the same in all species. First, the distribution of A-specific substitution rates in human gain regions were compared to that of stably marked regions. Although significantly elevated human-specific and ape-specific substitution rates were observed for human gain regions, this was also true for chimpanzee-specific and rhesus-specific substitution rates. An additional comparison of the values of the ratio of (human-specific substitution rate)/(chimpanzee-specific substitution rate) and (ape-specific substitution rate)/(rhesus-specific substitution rate) for human-gain regions versus stably-marked regions was conducted revealing no significant difference using a Wilcoxon rank-sum test.

Age analysis and evolutionary constraint

The 46-way MultiZ alignment was downloaded from the UCSC Genome Browser. Low quality genomes with less than 5x sequencing coverage were removed from the alignment, leaving 25 vertebrate genomes. Reproducibly marked promoters and enhancers were selected for age analysis. Sequence alignments of H3K27ac regions were extracted from the MultiZ alignments. Exons were subtracted from the alignment to remove their effects on the conservation of promoters. The percentage of alignable bases to the human reference was counted each species. The ages of the marked regions were estimated from the most distantly related species with at least 50% bases alignable to the human reference. In order to infer the evolutionary constraint of the promoters and enhancers only the 13 placental mammals from above alignment were considered. The phylogenetic tree for the 13 placental mammals was extracted from the existing 46-way phylogenic tree in the UCSC genome browser. An R implementation of the PHAST package was used to measure conservation (Hubisz et al., 2011). Neutral evolution rate was estimated across 1Mb windows of the genome using the REV model. Conservation within the indicated subtrees and the whole placental mammal tree was measured using the phyloP likelihood ratio test. The estimated divergence times of species were obtained from published studies (Bininda-Emonds et al., 2007; Hedges, 2002; Locke et al., 2011).

Supplementary References

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met *57*, 289–300.

Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., and Purvis, A. (2007). The delayed rise of present-day mammals. Nature *446*, 507–512.

Bullard, J.H., Purdom, E., Hansen, K.D., and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics *11*, 94.

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol *4*, P3.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Habegger, L., Sboner, A., Gianoulis, T.A., Rozowsky, J., Agarwal, A., Snyder, M., and Gerstein, M. (2011). RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. Bioinformatics *27*, 281–283.

Hedges, S.B. (2002). The origin and evolution of model organisms. Nat Rev Genet *3*, 838–849.

Hubisz, M.J., Pollard, K.S., and Siepel, A. (2011). PHAST and RPHAST: phylogenetic analysis with space/time models. Brief. Bioinformatics *12*, 41–51.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996–1006.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Locke, D.P., Hillier, L.W., Warren, W.C., Worley, K.C., Nazareth, L.V., Muzny, D.M., Yang, S.-P., Wang, Z., Chinwalla, A.T., Minx, P., et al. (2011). Comparative and demographic analysis of orang-utan genomes. Nature *469*, 529–533.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods *5*, 621–628.

Mu, X.J., Lu, Z.J., Kong, Y., Lam, H.Y.K., and Gerstein, M.B. (2011). Analysis of genomic variation in non-coding elements using population-scale sequencing data from the 1000 Genomes Project. Nucleic Acids Res *39*, 7058–7076.

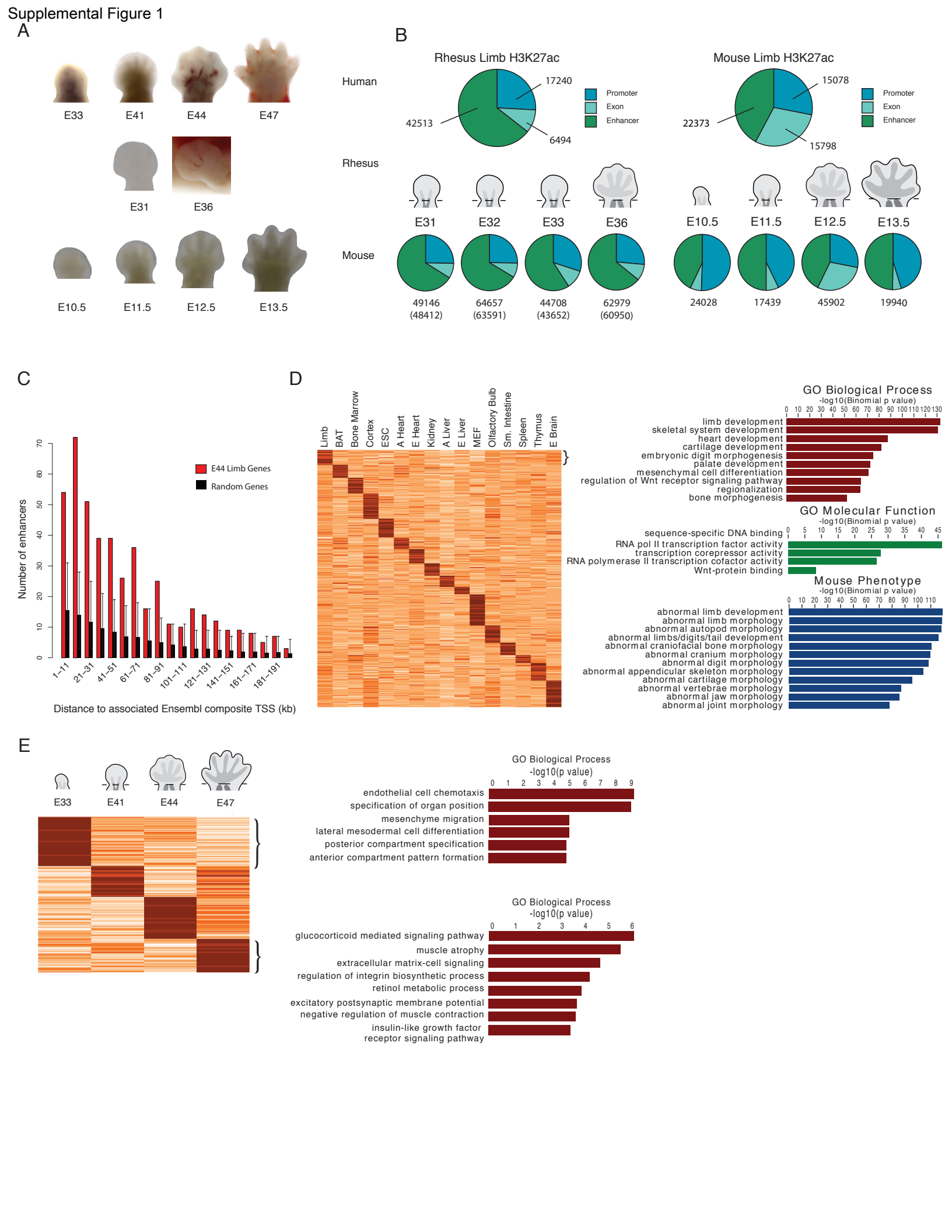
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744–752.

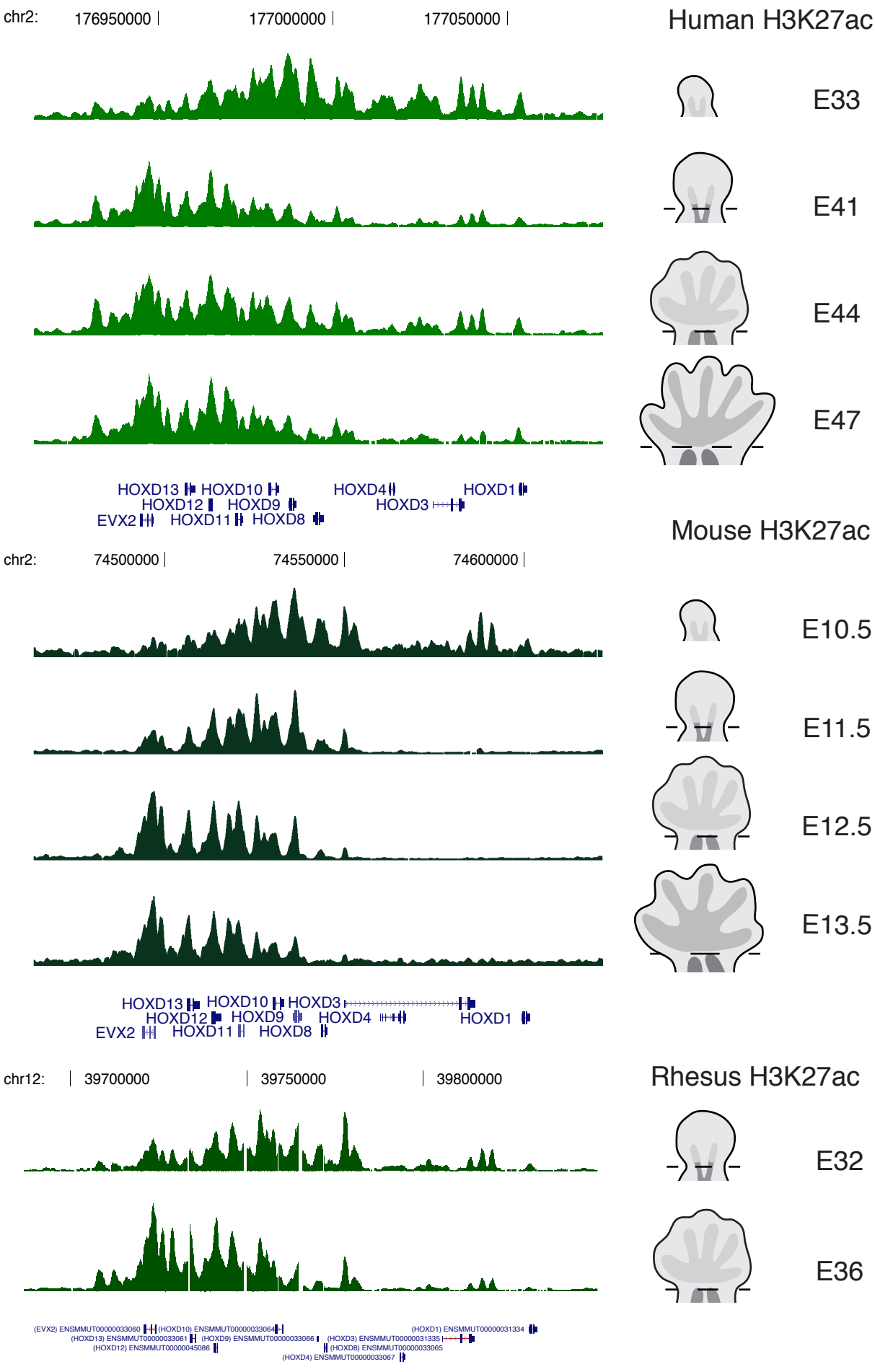
Sandelin, A., Alkema, W., Engström, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res *32*, D91–D94.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res *15*, 1034–1050.

Supplemental Figure 1

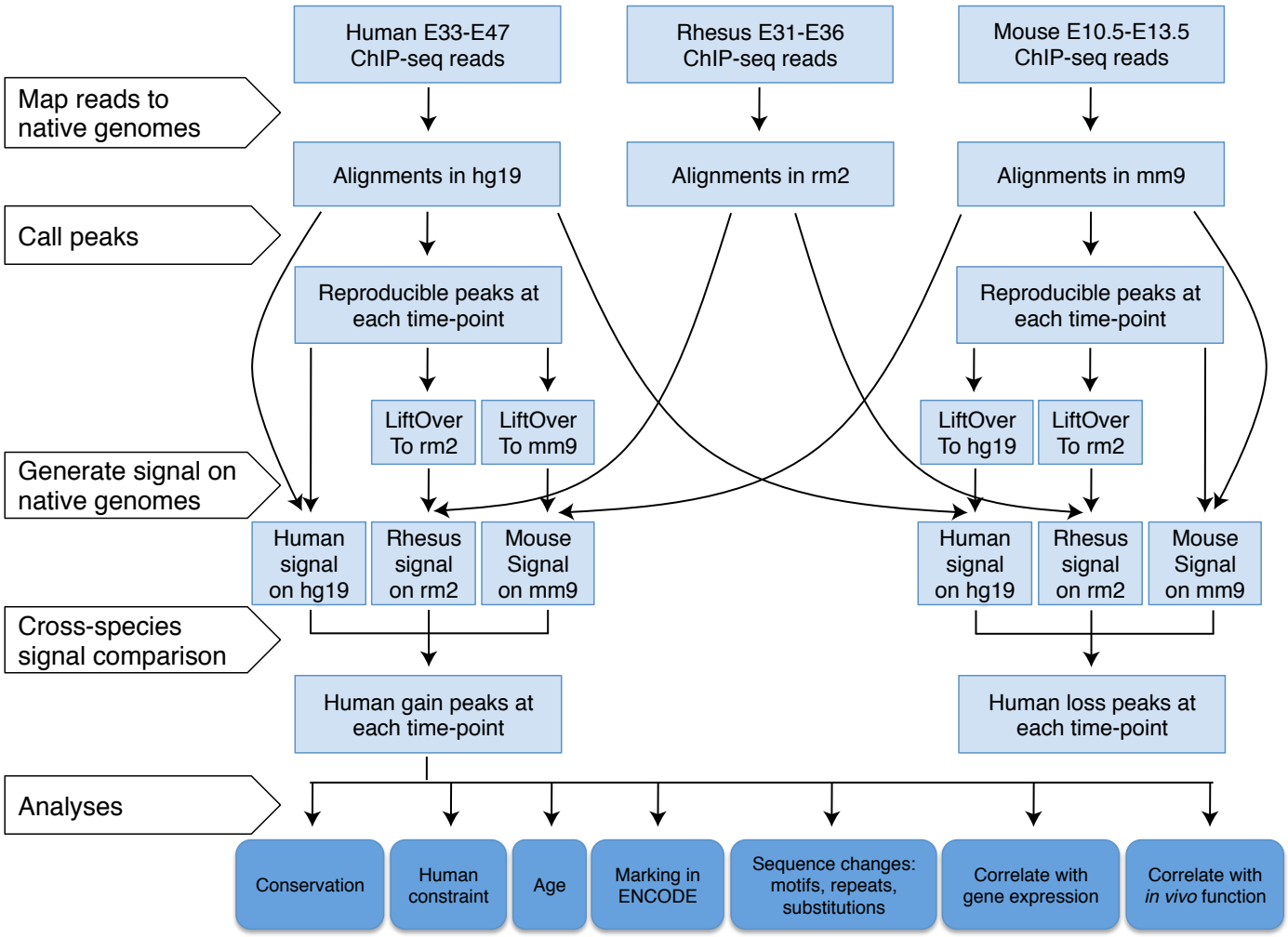


Supplemental Figure 2

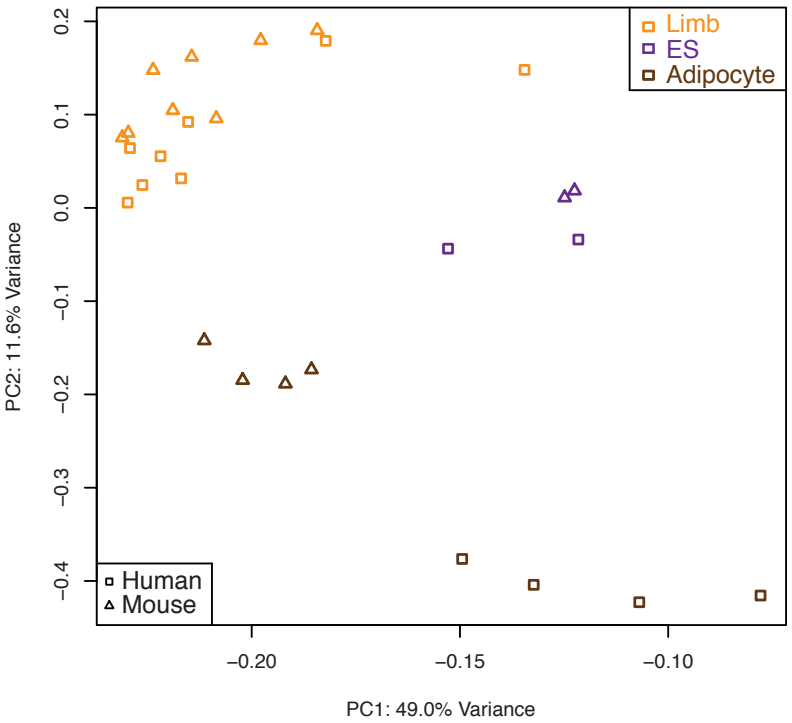


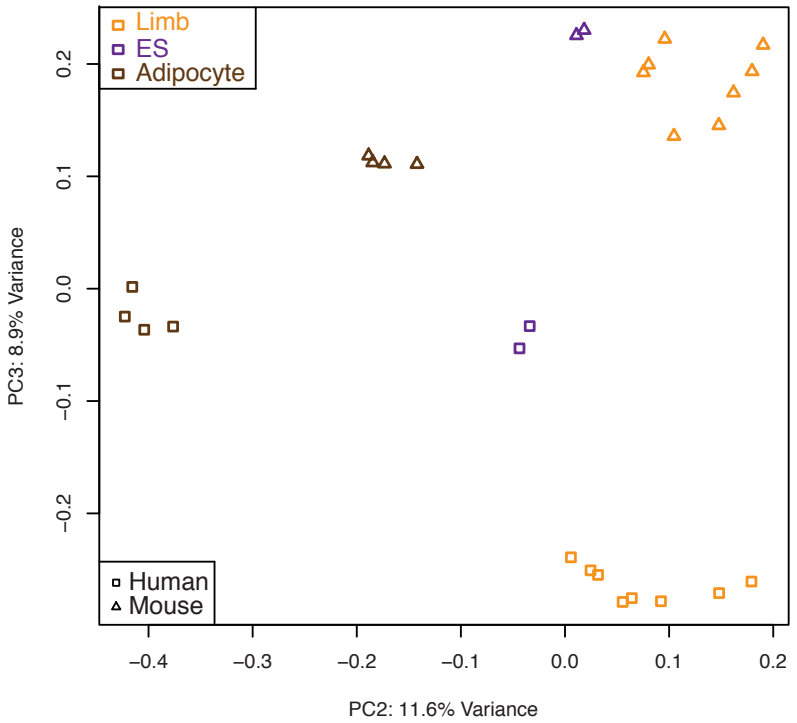
Supplemental Figure 3

A

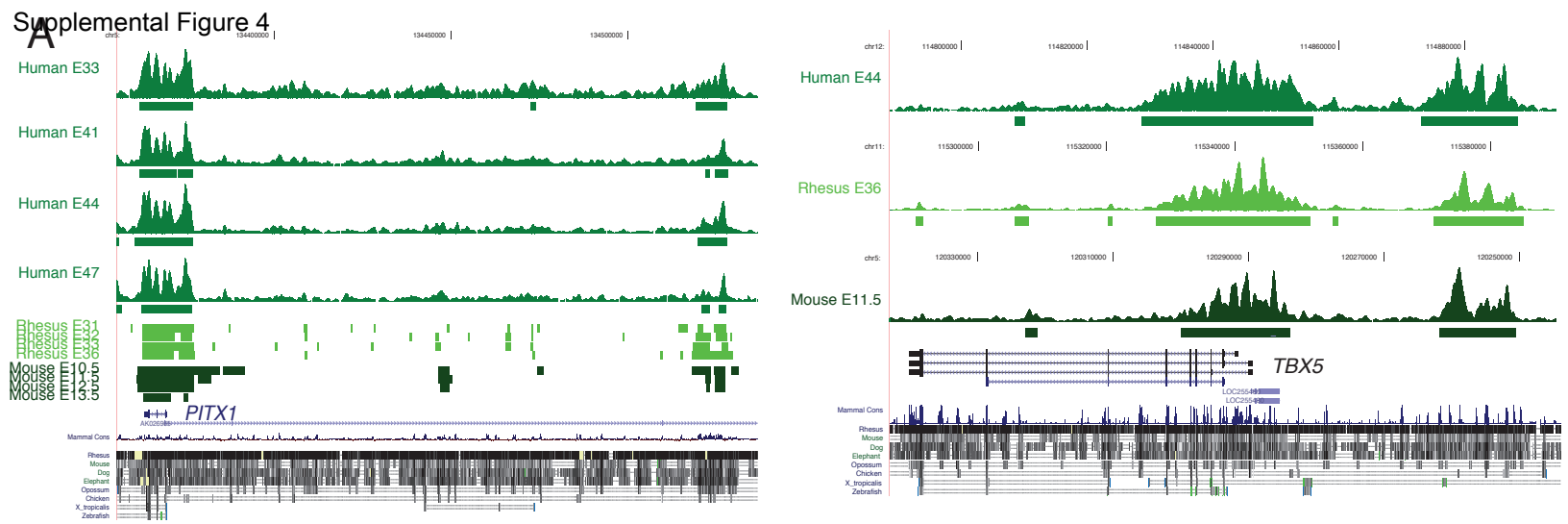


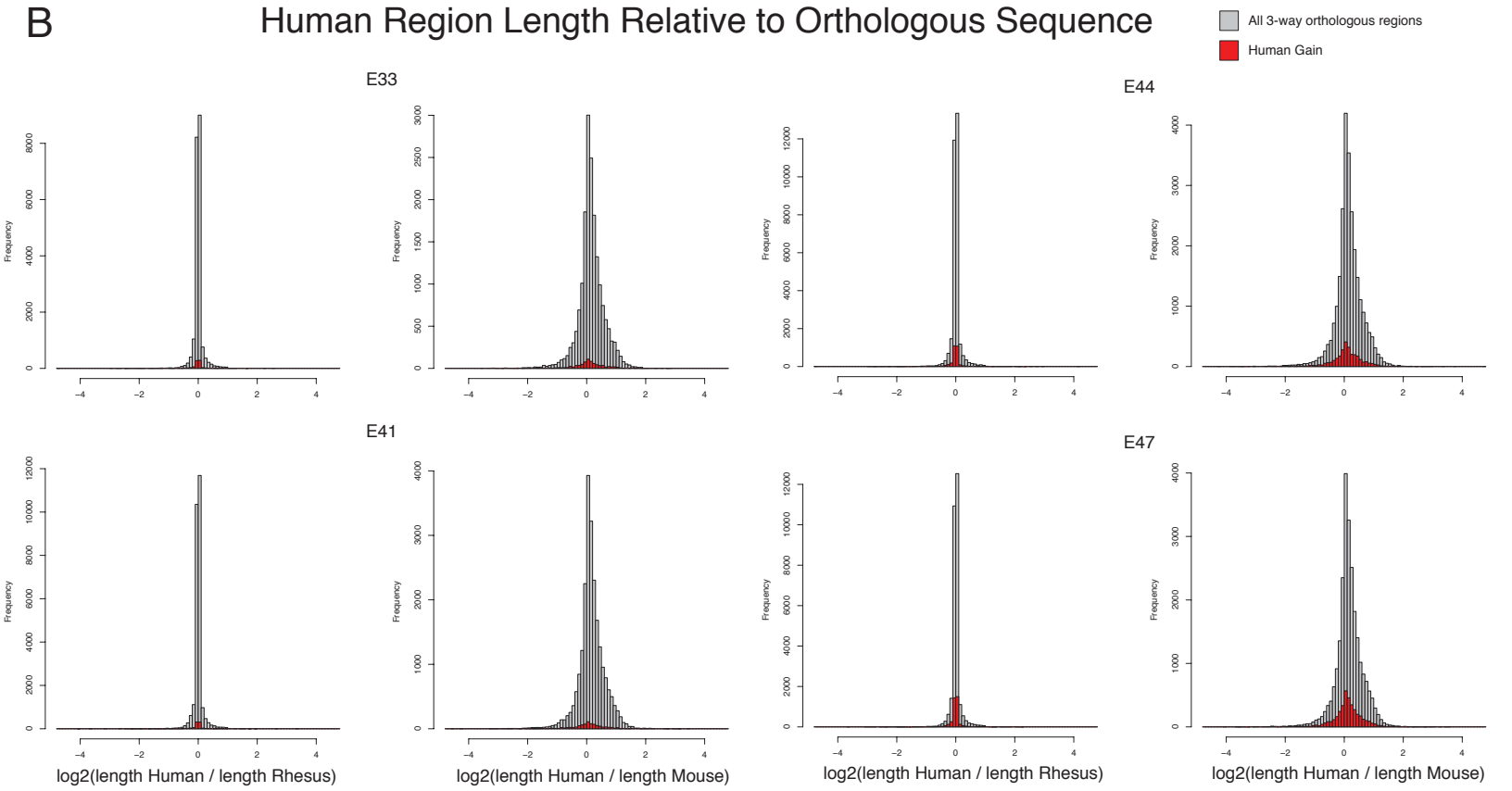
B

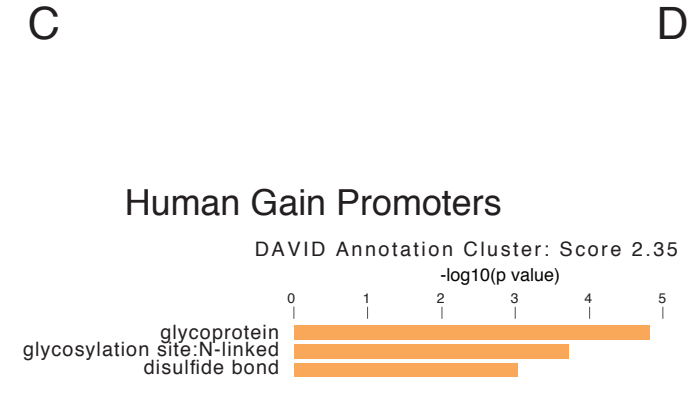


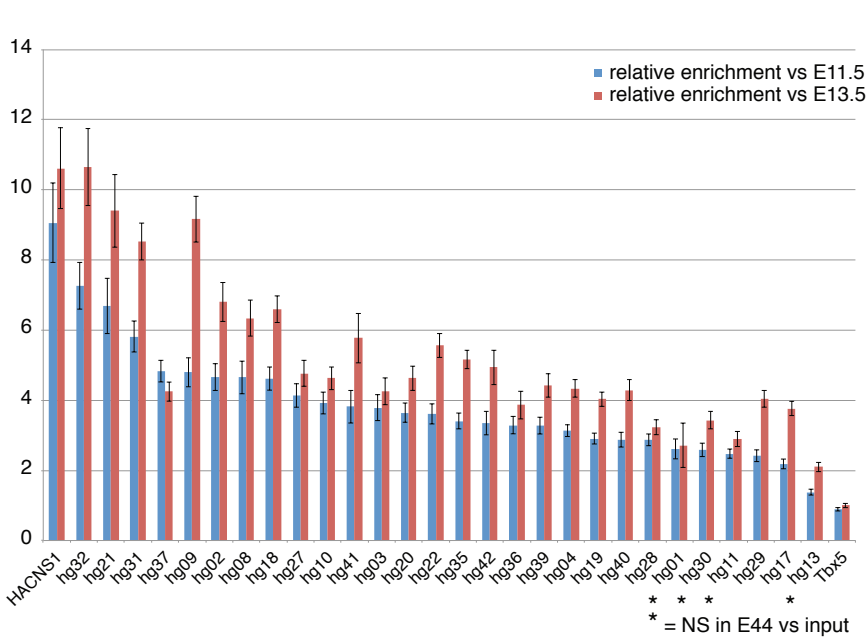


Supplemental Figure 4

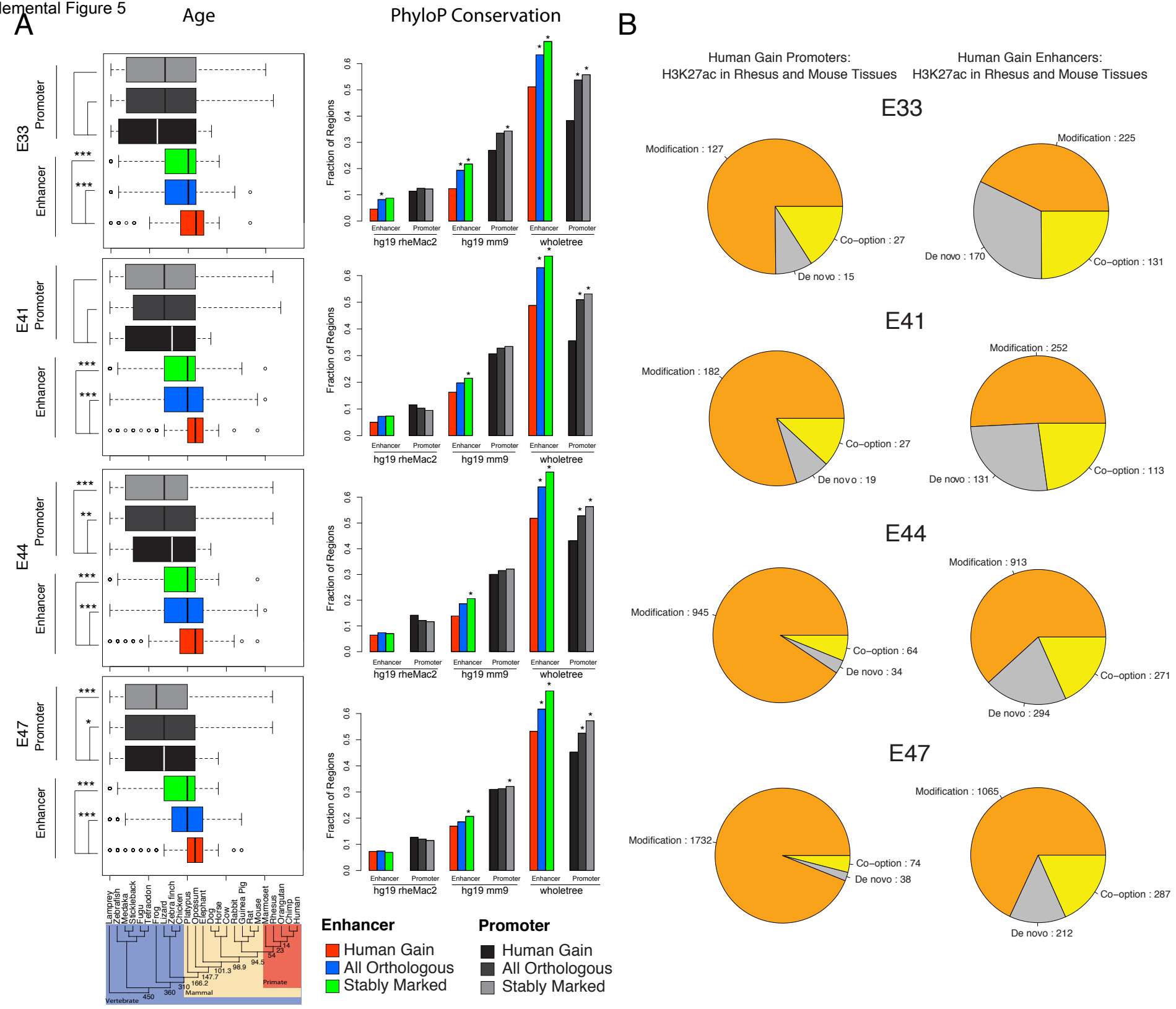






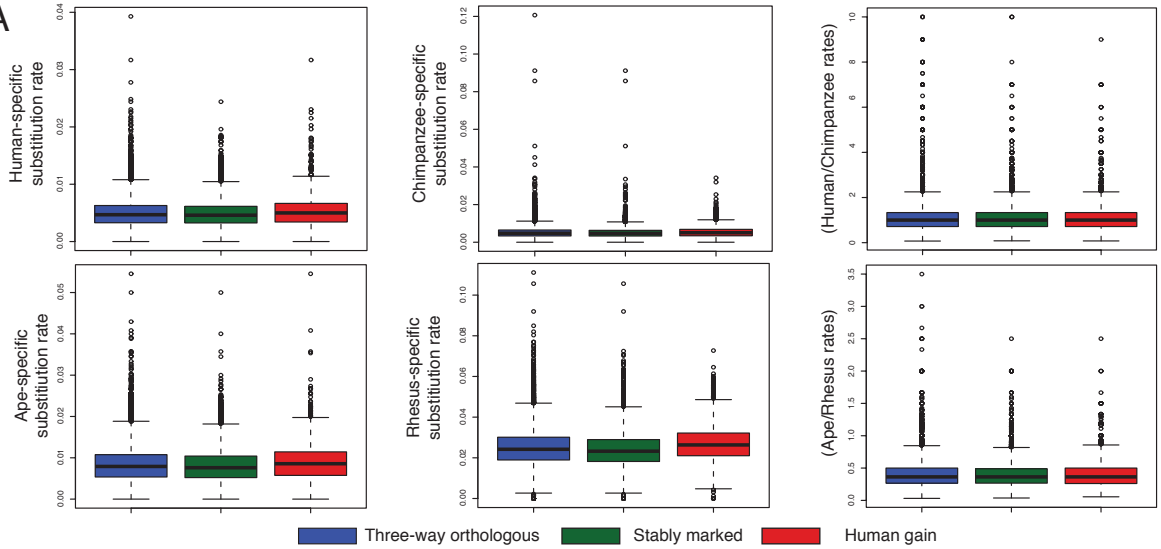


Supplemental Figure 5



Supplemental Figure 6

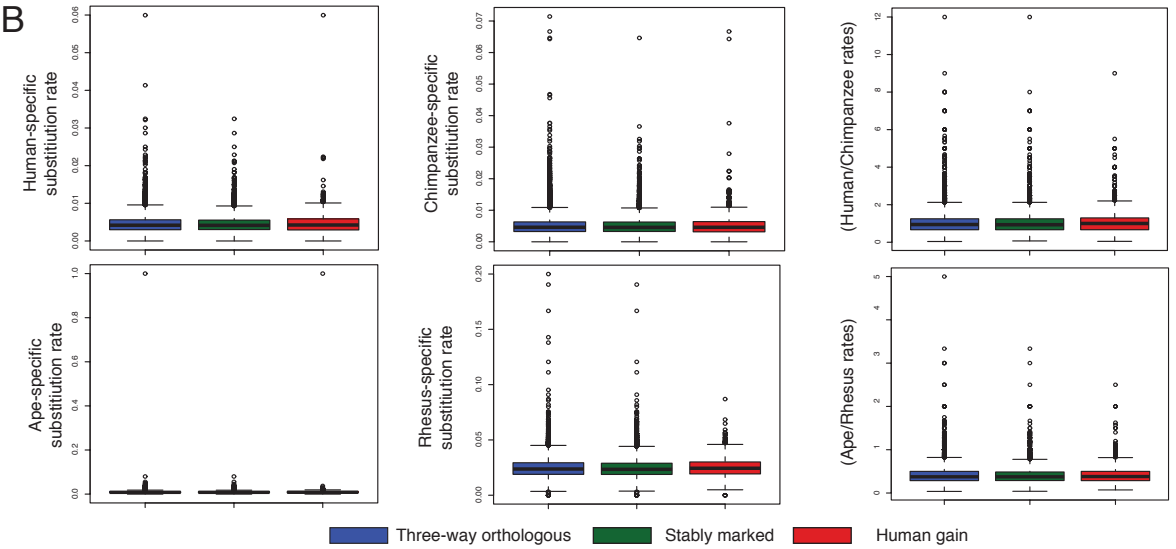
A



Enhancer	Human-specific Substitution rate	Chimpanzee-specific Substitution rate	Human/ Chimpanzee	Ape-specific Substitution rate	Rhesus-specific Substitution rate	Ape/Rhesus
E33	3.14e-04	3.26e-04	-6.65e-03	6.65e-04	3.73e-03	-2.62e-03
	0.120	0.018	0.296	0.018	7.54e-16	0.133
E41	6.46e-04	4.73e-04	-0.020	6.14e-04	3.71e-03	5.60e-03
	4.90e-05	2.97e-04	0.888	0.039	<2.2e-16	0.094
E44	4.40e-04	3.73e-04	0.024	8.84e-04	3.04e-03	0.016
	5.66e-06	7.10e-06	0.711	1.38e-10	<2.2e-16	0.448
E47	3.92e-04	2.62e-04	-0.017	9.54e-04	2.68e-03	0.013
	1.83e-06	2.64e-05	0.886	1.75e-10	<2.2e-16	0.727

difference (human gain - stably marked)
p-value

B

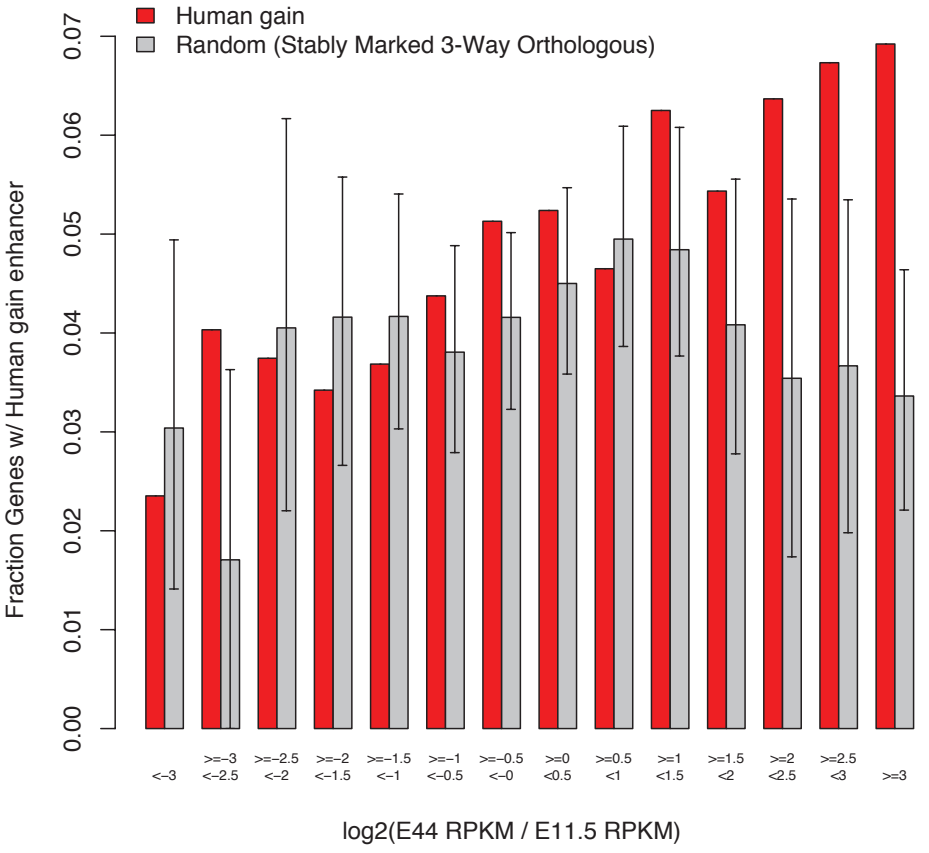


Promoter	Human-specific Substitution rate	Chimpanzee-specific Substitution rate	Chimpanzee	Ape-specific Substitution rate	Rhesus-specific Substitution rate	Ape/Rhesus
E33	5.56e-04	6.11e-04	0.047	7.79e-04	2.02e-03	0.024
	0.013	0.023	0.802	0.061	1.19e-04	0.764
E41	1.79e-04	-2.44e-04	0.074	1.12e-03	1.63e-03	0.042
	0.850	0.260	0.524	0.036	6.25e-03	0.254
E44	1.92e-04	1.29e-04	0.035	1.47e-03	7.91e-04	0.022
	0.316	0.855	0.166	3.71e-03	1.01e-03	0.108
E47	4.88e-05	7.06e-05	0.026	5.49e-04	1.22e-03	5.27e-03
	0.628	0.907	0.315	7.48e-07	7.71e-09	0.646

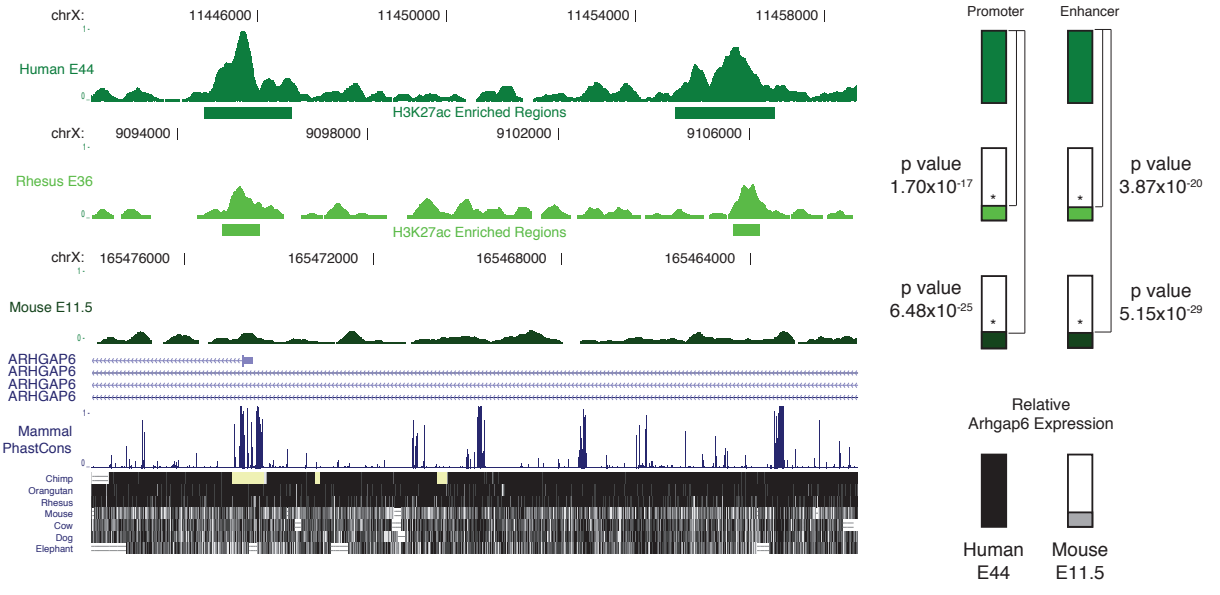
difference (human gain - stably marked)
p-value

Supplemental Figure 7

A



B



C