Functional Divergence of the Nuclear Receptor *NR2C1* as a Modulator of Pluripotentiality During Hominid Evolution

Jennifer L. Baker,[*,†,‡,1] Katherine A. Dunn,[§] Joseph Mingrone,[**] Bernard A. Wood,[*,††] Beverly A. Karpinski,[‡‡] Chet C. Sherwood,[*,†] Derek E. Wildman,[§§,*] Thomas M. Maynard,[†,†††] and Joseph P. Bielawski[§,**,1]**

*Center for the Advanced Study of Human Paleobiology, [†]George Washington Institute for Neuroscience, [‡‡]Department of Anatomy and Regenerative Biology, and [†††]Department of Pharmacology and Physiology, George Washington University, Washington, DC 20052, [‡]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, [§]Department of Biology and [**]Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2, [††]Human Origins Program, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, and [§§]Department of Molecular and Integrative Physiology and [***]Institute for Genomic Biology, University of Illinois Urbana-Champaign, Champaign, Illinois 61801

ORCID ID: 0000-0002-5192-2377 (J.L.B.)

ABSTRACT Genes encoding nuclear receptors (NRs) are attractive as candidates for investigating the evolution of gene regulation because they (1) have a direct effect on gene expression and (2) modulate many cellular processes that underlie development. We employed a three-phase investigation linking NR molecular evolution among primates with direct experimental assessment of NR function. Phase 1 was an analysis of NR domain evolution and the results were used to guide the design of phase 2, a codon-model-based survey for alterations of natural selection within the hominids. By using a series of reliability and robustness analyses we selected a single gene, *NR2C1*, as the best candidate for experimental assessment. We carried out assays to determine whether changes between the ancestral and extant *NR2C1*s could have impacted stem cell pluripotency (phase 3). We evaluated human, chimpanzee, and ancestral *NR2C1* for transcriptional modulation of *Oct4* and *Nanog* (key regulators of pluripotency and cell lineage commitment), promoter activity for *Pepck* (a proxy for differentiation in numerous cell types), and average size of embryological stem cell colonies (a proxy for the self-renewal capacity of pluripotent cells). Results supported the signal for alteration of natural selection identified in phase 2. We suggest that adaptive evolution of gene regulation has impacted several aspects of pluripotentiality within primates. Our study illustrates that the combination of targeted evolutionary surveys and experimental analysis is an effective strategy for investigating the evolution of gene regulation with respect to developmental phenotypes.

KEYWORDS ancestral gene reconstruction (AGR); codon models; hominid evolutionary survey; nuclear receptors; *NR2C1*; testicular receptor 2 (TR2); pluripotentiality

HUMAN evolutionary biology seeks to understand the origins of the defining characteristics of modern humans, such as our large brains, upright posture, obligatory bipedal gait, longevity, and extended juvenile period. While fossil morphology and artifacts recovered from archaeological sites are essential to inferring anatomical structure, function, and behavior in the past (Mcbrearty and Brooks 2000; Alemseged *et al.* 2006; Tryon *et al.* 2008; Jungers *et al.* 2009a,b; Braun *et al.* 2010; Ward *et al.* 2011), only through molecular genetic analyses can we make the ultimate connection between phenotype and genotype (Wood 1996; Allman *et al.* 2010; Boddy *et al.* 2012; Sherwood and Duka 2012). The eventual goal is to understand to what extent modern structures and functions are determined by different genetic systems and the extent to which the evolution of those systems has played a role in the evolution of the human lineage.

doi: 10.1534/genetics.115.183889
Manuscript received October 22, 2015; accepted for publication April 5, 2016; published Early Online April 11, 2016.
Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.183889/-/DC1.
[1]Corresponding author: Center for Research on Genomics and Global Health NHGRI/NIH 12 South Dr. Bethesda, MD 20892. E-mail: JenniferLBaker7@gmail.com; Department of Biology 1355 oxford Street Dalhousie University Halifax, Nova Scotia Canada B3H 4R2. E-mail: J.Bielawski@dal.ca.

A superfamily of transcription factors called the nuclear receptors (NRs) are attractive candidates for a combined evolutionary and functional investigation of hominids (*e.g.*, the clade that includes modern great apes and their last common ancestors). As transcription factors, NRs control many aspects of development, metabolism, reproduction, and endocrine signaling (Kohn *et al.* 2012). Their direct involvement in numerous physiological functions has motivated considerable research into their role in the evolution of hormone-mediated traits (Ketterson *et al.* 2009). Modern humans possess 48 NRs (Robinson-Rechavi *et al.* 2001) divided into six subfamilies known as NR 1–6 (Laudet 1997; Germain *et al.* 2006). Regulation of gene expression by NRs is typically induced by either endogenous or exongenous ligands, but there are some NRs, termed orphans, for which the ligand has yet to be identified or may not exist (Enmark and Gustafsson 1996; Benoit *et al.* 2006). Many functions of the NRs are sufficiently well characterized to permit the design of assays to investigate how amino acid changes could impact gene expression during key phases of primate development.

Because the NR family has been highly conserved for millions of years, an increase in the number of substitutions along a given lineage is often viewed as consistent with an underlying adaptive event. Such interpretations are premature, as amino acid changes also accumulate by neutral processes. In fact, there are only a handful of studies indicating that the dynamics of nonsynonymous evolution among primate NRs differ significantly from neutral expectations (Krasowski *et al.* 2005; Williamson *et al.* 2007; Chen *et al.* 2008). This is in part because sequence divergence among primates is characteristically low and it is challenging to differentiate adaptive substitutions from the background of neutral substitutions in such data. A full understanding of the phenotypic significance of primate NR evolution requires explicitly linking formal evolutionary analyses with experimental approaches focused on the functional impact of amino acid substitutions (Ugalde *et al.* 2004; Brayer *et al.* 2011; Kratzer *et al.* 2014).

Ancestral gene reconstruction (AGR), when combined with synthesis and assessment of the inferred ancestral protein in the laboratory, provides a powerful framework for investigating the contribution of extinct genetic systems to phenotypic change (Thornton 2001; Thornton *et al.* 2003; Bloch *et al.* 2015). Indeed, AGR has been used effectively to investigate the evolution of function in a wide variety of different proteins (Chang *et al.* 2002; Gaucher *et al.* 2003; Ugalde *et al.* 2004; Gaucher *et al.* 2008; Bridgham *et al.* 2009; Harms and Thornton 2010, 2013; Eick and Thornton 2011; Brayer *et al.* 2011). It is not feasible, however, to experimentally assess the functional consequences of even a subset of amino acids in each of the 48 primate NRs. Such a comprehensive survey would be prohibitively costly and time consuming. Molecular evolutionary modeling and analysis, on the other hand, is feasible on this scale and can provide information that can be used to identify candidate genes for further investigation (Chen *et al.* 2008). Here we describe an evolutionary analysis of NR sequence evolution, which identified *NR2C1* as a candidate for AGR-based experimental assays. Our functional assays revealed *NR2C1* as a potential modulator of pluripotentiality during hominid evolution.

Our investigation was divided into three phases. In the first phase, we employed fixed-effect (FE) codon models to characterize average rates and patterns of primate NR evolution with respect to its primary structural domains. Those results were used to guide the design of the second phase of analysis, which surveyed all 48 NRs for cases of hominid-specific alterations in the intensity of natural selection. We used a combination of branch-site codon models (developed to detect episodes of positive selection) and clade-site models (developed to detect any change in the distribution of selection pressures) to improve our capacity to detect functional divergence within NRs. Combining these models can uncover functionally relevant patterns of evolution that may not be apparent when they are used in isolation (Schott *et al.* 2014). Prior to interpreting the modeling results, we carried out a suite of reliability and robustness analyses, which led us to exclude a number of genes from further consideration. In the third phase, we selected a single gene (*NR2C1*) for further experimental investigation. We then inferred an ancestral amino acid sequence for *NR2C1*, generated an expression vector containing a synthetic open reading frame (ORF) that encodes this protein sequence, and tested *in vitro* whether evolutionary changes in its amino acid sequence impacted its function. Existing research on *NR2C1* suggests it is involved in neural differentiation and can act as an activator of the pluripotency factors *Oct4* and *Nanog* (Shyr *et al.* 2009). This information was used to design several *in vitro* assays, the results of which suggest that the function of the ancestral form of *NR2C1* differed from both the human and chimpanzee forms. We hypothesize that *NR2C1* may regulate aspects of the stem cell pluripotentiality, which could impact anatomical and physiological characteristics that distinguish humans from the other great apes.

Materials and Methods

Origin and processing of DNA sequences

The University of California Santa Cruz Genome Browser (Kent *et al.* 2002) contains 146 NR sequences for each of the 12 mammalian lineages (Figure 1) included in this study. We downloaded all the NR sequences (date of download: March 25, 2014) and obtained a provisional alignment for each using Multiple Alignment using Fast Fourier Transform (MAFFT) with default settings (Katoh 2002). These data were then filtered in a two-step process. First, all alignments were inspected for in-frame stop codons. These occurred within some alignments corresponding to splice variants at a single locus. We selected the alignment for the largest splice variant that did not contain an in-frame stop. This yielded one alignment per each human NR encoding gene

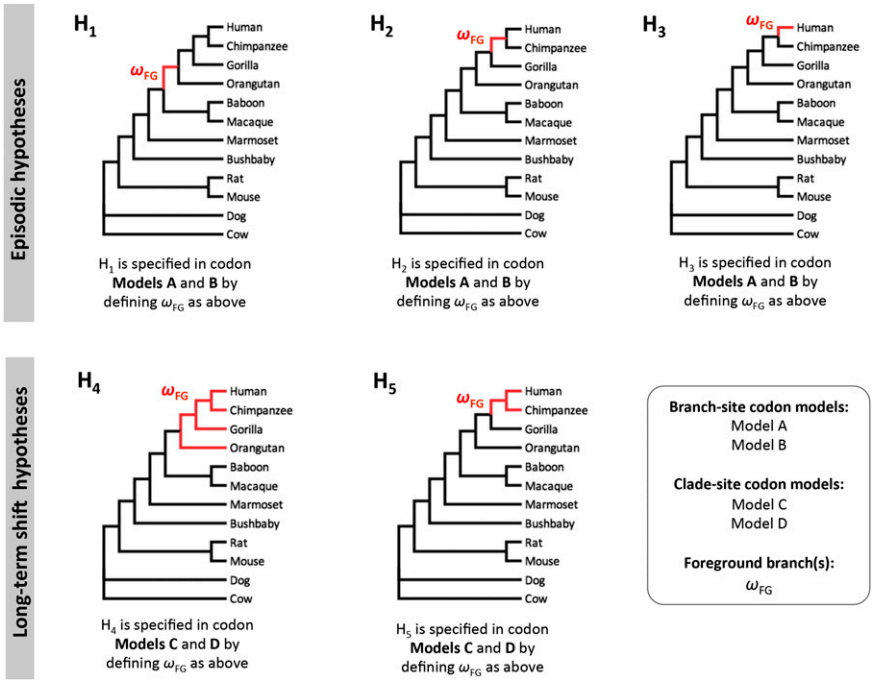

Figure 1 Phylogenetic relationships of the 12 mammalian lineages included in this study, and the alternative hypotheses for branch-site (LRT-1 and LRT-2) and clade analyses (LRT-3 and LRT-4). The red branches in hypotheses 1–3 are specified as the foreground branch in branch-site models A and B, which are employed to carry out LRT-1 and LRT-2 for episodic evolution. The red clades in hypotheses 4 and 5 are specified as the foreground clades in clade-site models C and D, which are employed to carry out LRT-3 and LRT-4 for long-term shifts in selection pressure.

(48 in total). Second, those 48 alignments were visually inspected and, where necessary, manually adjusted to improve the alignment or to exclude poorly aligned regions. Details for the original 146 sequences are provided in Supplemental Material, Table S1 and the final 48 alignments were deposited in the DRYAD data repository (datadryad.org: doi : 10.5061/dryad.bg3g3).

Assessing rates and patterns of evolution among NR structural domains

We investigated primate NR evolution with respect to four of the five major domains. We mapped each aligned site to (i) the N-terminal domain (NTD), (ii) the DNA binding domain (DBD), (iii) the flexible hinge domain (HD), and (iv) the ligand binding domain (LBD) using the Conserved Domain Database (Marchler-Bauer *et al.* 2011). The C-terminal domain (CTD) was excluded from this analysis, as it is not present in every NR (Bourguet *et al.*, 2000). This structural information was added to a codon model as a FE partition of an alignment, and the different partitions were allowed to have heterogeneous evolutionary dynamics (Yang and Swanson 2002; Bao *et al.* 2007). The structural partitions are included in the files deposited in the DRYAD data repository (doi : 10.5061/dryad.bg3g3). We fit the FE models to each alignment and employed likelihood ratio tests (LRTs) to test for heterogeneity among domains in selection intensity (ω), overall rate of evolution via a branch-length scale parameter (c), transition–transversion ratio (κ), and equilibrium codon frequencies (π_j) (Bao *et al.* 2007). The parameter ω serves as a measure of the intensity of natural selection pressure, with purifying, neutral, or positive selection indicated by values of < 1, $= 1$, or > 1 (Bielawski *et al.*, 2016). An expanded description of the approach is presented in File S1.

Surveying NRs for spatial and temporal variation in selective pressure

Detecting episodic evolution: We employed branch-site codon models A and B to test for sites experiencing a short-term (episodic) shift in the intensity of natural selection pressure (Yang and Nielsen 2002; Zhang *et al.* 2005). Models A and B permit the intensity of selection pressure to vary both among sites and among branches, with a model for episodic change obtained by specifying unique selection along a single branch (ω_{FG}) for a proportion of sites within a gene (p_{FG}). We used these models to test three *a priori* hypotheses about episodic evolution (ω_{FG} in Figure 1: H_1–H_3) at a fraction of sites, p_{FG}, within primate NRs. Models A and B are mixture models, and they employ additional parameters for sites where the intensity of selection is not episodic. Further details about the other parameters of the ω distribution are provided in File S2. We employed two LRTs for episodic evolution at a fraction of sites. LRT-1 compares constrained model A ($\omega_{FG} = 1$) to unconstrained model A ($\omega_{FG} > 1$). This LRT is intended as a formal test for an episode of positive selection (*i.e.*, a test for $\omega_{FG} > 1$). LRT-2 compares M3 ($p_{FG} = 0$) to model B ($p_{FG} > 0$). This LRT is employed to test for any episodic change in selection (*i.e.*, the episode need not involve $\omega_{FG} > 1$). Both LRTs were applied to each of the three episodic hypotheses shown in Figure 1 (H_1–H_3). Further details about these LRTs, and the involved codon models, are provided in File S2.

Detecting a long-term shift: We employed clade-site codon models C and D to test for a fraction of sites experiencing a long-term shift in the intensity of selection pressure (Yang and Nielsen 2002; Bielawski and Yang 2004). Models C and D permit the intensity of selection pressure to vary between entire clades, as well as among sites. A selective shift

is permitted by including a parameter for the proportion of sites (p_{SHIFT}) that have alternate ω's (ω_{FG} and ω_{BG}) between subtrees. As mixture models, they have additional parameters for the intensity of selection at other sites (and selection is assumed to be homogenous over the tree at those other sites). Further details about the ω distributions for these models are provided in File S2. We used models C and D to test *a priori* hypotheses about clade-level selective shifts in primates (Figure 1, H_4 and H_5) within NR genes. We employed two different LRTs (LRT-3 and LRT-4) to test those hypotheses. LRT-3 compares model M2a-rel to clade-site model C (Weadick and Chang 2012). This LRT-3 is intended as a test for sites having positive selection across an entire clade (*i.e.*, $p_{SHIFT} > 0$ and $\omega_{FG} > 1$ or $\omega_{BG} > 1$). LRT-4 compares model M3 to clade-site model D (Bielawski and Yang 2004). LRT-4 is employed as a generalized test for any type of long-term shift in the intensity of selection (*i.e.*, a test for $p_{SHIFT} > 0$, regardless of the values of ω_{FG} and ω_{BG}). Further details about these LRTs, and the involved codon models, are provided in File S2. The codeml program from version 4.4 of the PAML package (Yang 2007) was used for likelihood calculation and parameter estimation under branch-site and clade-site codon models.

Assessing the quality of the signal for variation in selective pressure through a suite of reliability and robustness analyses

We inferred the subgroups of genes related by a given evolutionary event (H_1–H_5) by controlling the false discovery rate (FDR) within each subgroup according to the method of Storey (2002). For every gene within a subgroup, we carried out a series of additional analyses to assess the reliability of the signal and robustness to model assumptions. First, because incorrect positional homology can impact some inferences (Schneider *et al.* 2009; Fletcher and Yang 2010), two coauthors independently assessed each alignment. Second, because our analyses were carried out under the species tree, we also estimated gene trees with RAxML (Stamatakis 2014) and reanalyzed the data under that topology. Third, we reanalyzed the data assuming substitution probabilities are proportional to the equilibrium value of the target nucleotide (MG94-style codon model: Muse and Gaut 1994) rather than the equilibrium value of the target codon (GY94-style codon model: Goldman and Yang 1994). Fourth, we used the Genetic Algorithm for Recombination Detection-Multiple Breakpoint (GARD-MBP) method (Kosakovsky Pond *et al.* 2006) to test for evidence of within-gene recombination events, as these can negatively impact some LRTs (Anisimova *et al.* 2003). Fifth, we employed the multilayer codon model of Rubinstein *et al.* (2011) to investigate whether site variability in the baseline rate of DNA/RNA substitution (*e.g.*, due to synonymous rate variability) impacted our initial estimates of ω.

Lastly, we employed nonparametric bootstrapping to quantify the uncertainty in the estimates obtained for the parameters of the ω distribution and to assess it for signs that statistical regularity conditions might not have been met (*e.g.*, bimodal distributions). Bootstrapping is a procedure of "sampling with replacement" that is routinely employed to assess phylogenetic inference, but has only recently been applied to ω-based inference of selection pressure (Bielawski *et al.* 2016). In this procedure, site patterns within the original multisequence alignment are sampled at random (with replacement) to create many new alignments with different distributions of site patterns. To investigate the properties of the ω distribution, all model parameters (including branch lengths) are reestimated from each bootstrap dataset. We used 100 bootstrap datasets to approximate the maximum likelihood estimate (MLE) distribution for the p_i and ω_i parameters of a codon model. These data were used to obtain 95% C.I.s for those parameters. We also used these distributions to determine when standard regularity conditions might not have been met. Reliable interpretation of the LRT assumes that they have been met, but discretization of continuous ω distributions within the codon models can make this difficult for some datasets (Bielawski *et al.* 2016). Violation of regularity conditions can be diagnosed as nonstandard MLE behavior such as strongly bimodal distributions for the p_i and ω_i parameters.

Ancestral sequence reconstruction

Reconstruction of an ancestral sequence was carried out for both DNA and amino acid states using PAML 4.4 (Yang 2007). The DNA-based reconstructions were based on the general time reversible (GTR) model (Yang 1994), and the amino-acid-based reconstructions were based on the Whelan and Goldman (WAG) replacement matrix (Whelan and Goldman 2001). Additional details, including the reconstructed states and their posterior probabilities, are provided in File S3. These analyses generated almost identical ancestral sequences, with the exception of one amino acid. The ancestral sequence inferred using the amino acid model had higher posterior probabilities for each state, and thus was chosen for gene synthesis. The full ancestral sequences for the human, chimpanzee, and the inferred last common ancestor (LCA) are presented in File S3.

Sequencing of NR2C1 transcripts

We first used the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) to search for polymorphisms. After confirming that the between-species differences were not associated with polymorphisms within either humans or chimpanzees (File S4), we partially sequenced complementary DNA (cDNA) samples prepared from cortex samples. The samples consisted of frozen postmortem brain tissue from the cortex of 11 nonhuman primate individuals. The fresh frozen brains were obtained from a variety of biomedical and zoological institutions. All animals were housed in accordance with National Institutes of Health, US Department of Agriculture, and the Animal Welfare Act regulations, and overseen by the Institutional Animal Care and Use Committees (IACUCs) of the respective institutions (and approved by George Washington University IACUC protocol A117). The time between death

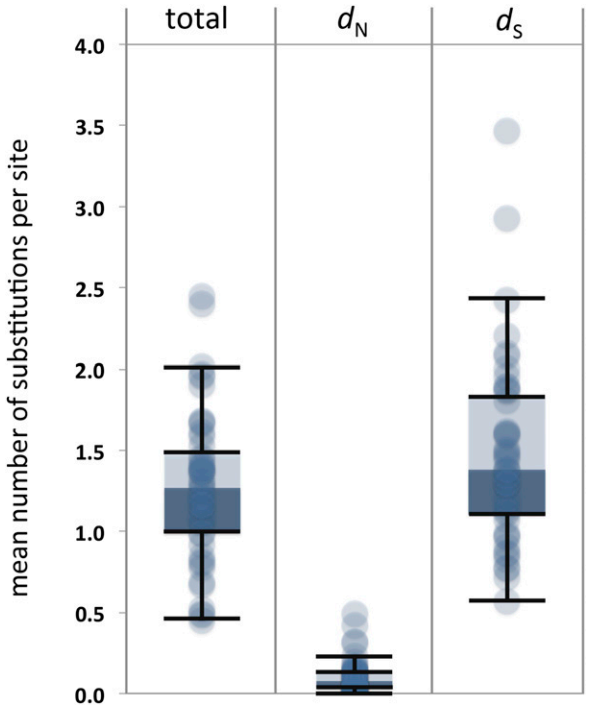

Figure 2 Distribution of maximum likelihood estimates of total tree length (t), the nonsynonymous rate (d_N), and the synonymous rate (d_S) among the 48 NR gene sequences. Sequence divergence statistics are derived from maximum likelihood estimates of parameter values under codon model M0. The scale for the total tree length is the number of substitutions per codon site, which consists of three nucleotides. The scale for both the nonsynonymous and synonymous rates is the number of substitutions per single nucleotide site of the relevant type. The distributions are summarized as box and whisker plots. Data points outside the top and bottom fences are interpreted as outliers.

and tissue freezing was, for those cases in which this information was available, never longer than 24 hr. Frozen brain samples were stored at −80° until use. No neurological deficits were detected in any of the individuals included in this study, and all brains appeared normal on routine inspection at necropsy.

RNA was extracted and complementary DNA (cDNA) prepared as per Maynard *et al.* (2013); briefly, RNA was extracted by Trizol extraction (Invitrogen), and contaminating genomic DNA was removed by DNAse digestion (DNAfree Turbo, Ambion). cDNA was generated by random-hexamer-primed first-strand cDNA synthesis, using ImPromII Reverse Transcriptase (Promega). We PCR amplified four independent amplicons that overlapped the key polymorphisms identified in the AGR analysis. Amplification of each cDNA was performed using standard Taq polymerase (Qiagen), with PCR primers listed in Table S2. All primers were designed to regions where genomic sequencing shows complete conservation of the DNA sequence. Each PCR product was gel purified, quantified, and sent for sequencing using both the forward and reverse primers used for the initial PCR. Sequences were aligned using Sequencher software and each chromatogram was manually inspected to validate each point where sequences were divergent. To further validate that the amino acid substitutions were unique to extant humans and chimpanzees, we sequenced amplicons from additional primate samples from the following species: *Pan troglodytes* (3), *Symphalangus syndactylus* (1), *Papio anubis* (1), *Macaca mulatta* (3), *Macaca nemestrina* (2), and *Pithecia pithecia* (1). Additional details are provided in File S4. As a reference, and for comparison, we also sequenced a human *NR2C1* coding sequence from a commercially available plasmid clone containing the *NR2C1* ORF in the shuttle vector pFN21A (Promega). Accession nos. for the novel *NR2C1* sequences generated in this study are: KT032104–KT032114.

Creation of expression vectors

We generated plasmid expression vectors containing the ORF of the human, chimpanzee, and inferred ancestral sequence reconstruction of *NR2C1*, by cloning these vectors into the pCINeo4 plasmid, which contains the composite cytomegalovirus (CMV)/chicken β-actin enhancer/promoter from pCAGG, the multicloning site and encephalomyocarditis virus–internal ribosomal entry site (EMCV–IRES) sequence from pIRES2–EGFP (Clontech), and a neomycin coding frame. The human ORF was cloned into the *Sal*I–*Bam*HI sites of this vector by PCR amplifying the ORF of the human *NR2C1* clone described above, with a forward primer that adds a *Sal*I site and canonical Kozak's consensus site (GTCGACCACC) at the 5′ end of the ORF, immediately preceding the ATG start codon, and adding a *Bam*HI site to the 3′ end, in frame with a human influenza hemagglutinin (HA) epitope tag in the vector. Chimpanzee and ancestral sequence expression vectors were generated by creating human codon-optimized sequences corresponding to their respective amino acid sequences, with similar *Sal*I/Kozak's consensus sites and *Bam*HI sites added to the 5′ and 3′ ends, respectively. These DNA fragments were synthesized (GeneBlocks, Integrated DNA Technologies) and cloned in a similar fashion into the *Sal*I and *Bam*HI sites of pCINeo. A small interfering RNA (siRNA) knockdown vector was generated by cloning a synthetic oligonucleotide into a short-hairpin RNA expression vector derived from pSilencer (Invitrogen). This siRNA cassette was then subcloned into the backbone of an expression vector (pSCH) containing the composite CMV/chicken β-actin (pCAGG) enhancer/promoter fused to a mCherry–IRES–hygromycin cassette. Control plasmids containing the pCINeo plasmid without the *NR2C1* insert, and containing the pSCH siRNA expression plasmid with a nonsense-sequence insert, were also generated. All plasmids were fully sequence verified before use. Endotoxin-free DNA preparations were made of each vector for use in embryonic stem (ES) cell experiments (EZNA Plasmid Maxi Kit, Omega Biotek).

ES cell cultures

Mouse ES cells [E14Tg2a, American Type Culture Collection (ATCC)] were cultured on a feeder layer in ES cell media containing Dulbecco's minimum essential medium (Invitrogen), supplemented with 15% fetal calf serum (HyClone),

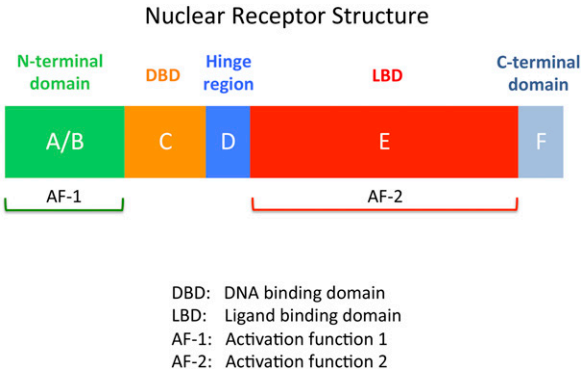

Figure 3 The structural domains of nuclear receptors (NRs). Sites within NRs are typically classified into six regions (called regions A–F), which correspond to five structural domains. The A/B region contains activation function 1, and corresponds to the N-terminal domain (NTD) of the protein. Region C is the highly conserved DNA binding domain (DBD). Region D is a flexible hinge region that connects the DBD to region E. Region E corresponds to the ligand binding domain (LBD) and contains activation function 2. The C-terminal domain (CTD, region F) varies in length between the different NRs and is nonexistent in some NRs.

0.1 mM 2-mercaptoethanol, penicillin/streptomycin/amphotercin-B (Anti-Anti, Invitrogen), and 100 units/ml leukemia inhibitory factor (LIF) (Enzo). Feeder layers were generated from confluent cultures of STO fibroblast cells (ATCC) by mitotically inactivating the cells with a 2-hr treatment with 10 μg/ml mitomycin C and passaging the cells at a 1:2 ratio onto gelatin-coated tissue culture plastic dishes. Stock ES cells were routinely maintained by trypsinizing and passaging confluent plates of ES cells at a 1:4 subculture ratio every 3–4 days.

Transfection of ES cells and clonal cultures

To transfect ES cells with expression vector plasmids, confluent plates of ES cells were trypsinized, washed twice in PBS, and resuspended in OptiMem (Invitrogen). The 10×10^6 ES cells were resuspended in 800 μl aliquots, and placed in a 4-mm cuvette, along with 5 μg each of two separate plasmids (a pCINeo plasmid expressing a variant of *NR2C1* or control, and a pSCH plasmid expressing the siRNA knockdown or the nonsense control). A total of five plasmid conditions were assayed and annotated as follows: nonsense (nonsense siRNA plasmid + control pCINeo plasmid), knockdown (knockdown siRNA plasmid + control pCINeo plasmid), human (knockdown siRNA plasmid + pCINeo containing *hNR2C1* ORF), chimpanzee (knockdown siRNA plasmid + pCINeo containing *cNR2C1* ORF), and ancestral (knockdown siRNA plasmid + pCINeo containing the inferred *aNR2C1* ORF). The cells were electroporated with two pulses of 500 V for 1 ms using a square-pulse electroporator (BTX). Electroporated cells were plated onto feeder layers with LIF containing media in six-well plates. After 24 hr, media were changed to fresh media containing both 300 μg/ml hygromycin and 300 μg/ml G418, to select for ES cells cotransfected with both expression vectors. After 18–21 days, single selected clones were apparent on the ES cell plates; we selected single clones, dissociated

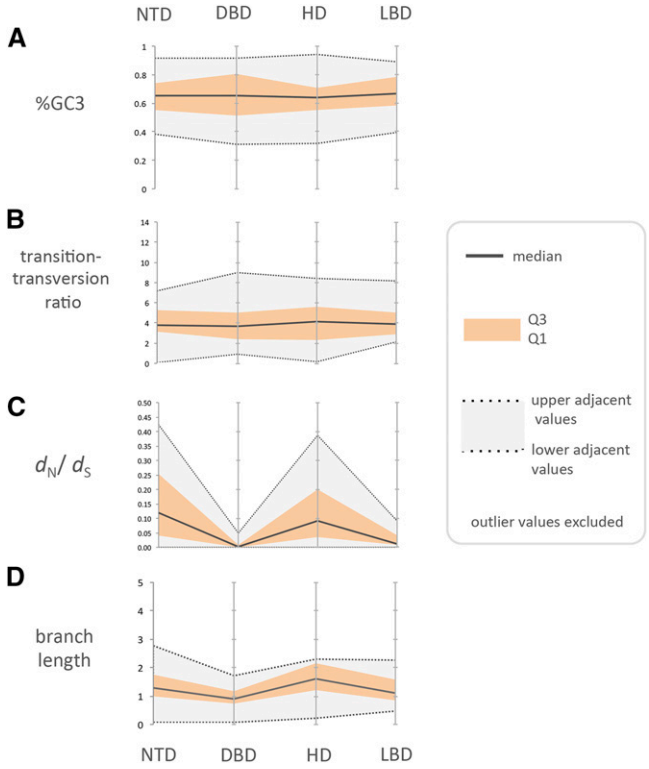

Figure 4 Summary of the variability in the evolutionary process among domains of NR proteins as inferred from codon models where the domain structure was specified as a fixed effect in the model. Data are from 46 NR genes partitioned into four categories: N-terminal domain (NTD); DNA binding domain (DBD); flexible hinge domain (HD); and the ligand binding domain (LBD). Codon model FE1 was used to obtain partition-specific estimates (A) transition–transversion ratio (κ); (B) equilibrium codon frequencies, which are approximated by %GC3 in this figure; (C) selection intensity ($\omega = d_N/d_S$); and (D) the overall rate of evolution (via a branch-length scale parameter, c). Each figure summarizes the distribution of the estimates over 46 NR genes. The dotted lines on the boundary of the gray band are equivalent to the top and bottom fences in a box and whisker plot [*i.e.*, Q3 + (1.5 × interquartile range; IQR) and Q1 − (1.5 × IQR)], and are referred to as the upper and lower adjacent values. Values beyond the upper and lower adjacent values in the plot are considered outliers and for clarity are not displayed in these plots.

and counted the total cell numbers, and plated 2000 cells/well into individual wells of six-well plates, in the same selection media (ES media with hygromycin and G418). These clonal ES cell cultures were repropagated every 18–21 days by selecting 10 clones (chosen from a set of contiguous clones in a single field to minimize any potential selection bias), and then this set of cells was again dissociated in trypsin, counted on a hemocytometer, and replated at 2000 cells/well to facilitate analysis over multiple generations.

Quantitative real-time PCR analysis of expression

Trizol-extracted RNA was obtained from pools of ES cells grown during the first passage from six independent clonal lines for each of the five plasmid conditions. cDNA was generated as described above and quantitative real-time PCR (qPCR) was performed to assess expression levels for mouse *NR2C1*, *Oct4*, and *Nanog* using the primers listed in

Table 1 Significant LRTs for a subset of sites having experienced an episodic alteration of selection pressure

Gene	$2\delta l$	P	q	p_i	ω_i
H_1: Great Ape					
LRT-1: model A (ω_{FG} = 1) *vs.* model A (ω_{FG} > 1)					
None	N.A.	N.A.	N.A.	N.A.	N.A.
LRT-2: M3(k = 2) *vs.* model B					
NR0B2	7.07	0.0292	0.7012	$P_{FG(a+b)}$ = (0.35 + 0.65)	ω_{FG} = 0.0
				p_1 = 0.00	ω_1 = 0.59
				p_0 = 0.00	ω_0 = 0.05
H_2: Human–Chimpanzee					
LRT-1: model A (ω_{FG} = 1) *vs.* model A (ω_{FG} > 1)					
None	N.A.	N.A.	N.A.	N.A.	N.A.
LRT-2: M3(k = 2) *vs.* model B					
None	N.A.	N.A.	N.A.	N.A.	N.A.
H_3: Human					
LRT-1: model A (ω_{FG} = 1) *vs.* model A (ω_{FG} > 1)					
NR1D1	10.27	0.0013	0.0650	$P_{FG(a+b)}$ = 0.002	ω_{FG} = 99
				p_1 = 0.05	[ω_1 = 1]
				p_0 = 0.95	ω_0 = 0.04
LRT-2: M3 (k = 2) *vs.* model B					
NR1D1	**15.25**	**0.0005**	**0.0234**	$P_{FG(a+b)}$ = 0.01	ω_{FG} = 99
				P_1 = 0.05	ω_1 = 0.86
				p_0 = 0.94	ω_0 = 0.04
PPARG	9.71	0.0077	0.1867	$P_{FG(a+b)}$ = 0.01	ω_{FG} = 35
				p_1 = 0.11	ω_1 = 0.46
				p_0 = 0.88	ω_0 = 0.0
NR2C1	8.70	0.0129	0.2063	$P_{FG(a+b)}$ = (0.24 + 0.76)	ω_{FG} = 99
				p_1 = 0.00	ω_1 = 0.33
				p_0 = 0.00	ω_0 = 0.02
PGR	7.67	0.0215	0.2586	$P_{FG(a+b)}$ = (0.12+0.05)	ω_{FG} = 6.23
				p_1 = 0.24	ω_1 = 0.58
				p_0 = 0.59	ω_0 = 0.05

Genes having a q-value of <0.05 are shown in boldface type. The foreground (FG) branches are fully specified for each hypothesis in Figure 2. The null model for all LRTs assumes homogenous selection pressure for all branches ($\omega_{BG} = \omega_{FG}$). LRT-1 has d.f. = 1. LRT-2 has d.f. = 2. The q-value is the expected proportion of false discoveries expected if the single-test P-value is used as the boundary to control the FDR. The parameter $P_{FG(a+b)}$ represents the proportion of sites subject to a change in selection intensity.

Table S2. Expression was normalized using *Gapdh* primers as an internal control (Table S2). Reactions were assembled using an EpMotion 5070 liquid handling system (Eppendorf) that combines forward and reverse gene-specific primers, with 7.5 μl of SsoFast EvaGreen Supermix (Bio-Rad) in a 14-μl reaction. qPCR analysis was performed using a CFX-384 Real-Time PCR Detection System.

Promoter activation assay

To assess the transcriptional activity of the *NR2C1* variants, we created a luciferase reporter vector using a synthetically generated 465-bp fragment of the mouse *Pepck* promoter (Roesler *et al.* 1989). HEK-293 cells (ATCC) were plated in 24-well plates in standard cell culture media (DMEM with 10% FCS and antibiotics) at 50% confluency and cotransfected with a reporter plasmid and an *NR2C1* variant using Effectene reagent (Qiagen) according to the manufacturer's protocol. After overnight incubation, luciferase activity was measured using a modified luciferase reporter assay system (Gold Biotechnology), and normalized to cells cotransfected with the reporter and an empty expression vector. Each cotransfection was repeated six times, with each assayed in duplicate.

Data availability

DRYAD data repository for alignments and structural partitions is as follows: doi : 10.5061/dryad.bg3g3. GenBank accession nos. for the novel *NR2C1* sequences generated in this study are as follows: KT032104–KT032114. The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

Results and Discussion

Phase 1: Characterizing NR domain evolution

We assessed sequence evolution in the NR family as represented by a set of 12 mammalian lineages (Figure 1). The data were composed of 48 NR alignments. Only moderate levels of sequence divergence were observed. The median tree length (across the 48 NR genes) was just 1.27 substitutions per codon, and partitioning total sequence divergence into synonymous and nonsynonymous components revealed that the majority of change was synonymous (Figure 2). The low levels of sequence divergence, and the sparseness of nonsynonymous change in particular, raise the possibility

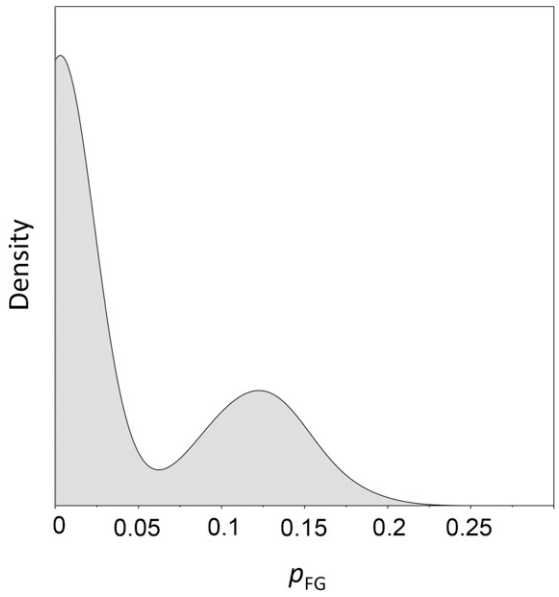

Figure 5 The sampling distribution for the proportion of sites evolving under episodic evolution (p_{FG}) for *NR1D1*. The sampling distribution was estimated from 100 replications of the nonparametric bootstrap under branch-site model B. Because the estimated distribution for p_{FG} is strongly bimodal, with considerable density at zero, the conditions necessary for reliable MLEs and LRTs cannot be assumed for *NR1D1*.

Table 2 Significant LRTs for a subset of sites having experienced a shift in the intensity of selection pressure at the origin of the great apes (H_4)

Gene	$2\delta l$	P	q	p_{SHIFT}	ω_{BG}	ω_{FG}
LRT-3: M2a-rel *vs.* model C						
RORA	**47.02**	**7.04e-12**	**1.55E-10**	**0.11**	**0.26**	**3.40**
RARG	**15.88**	**6.76e-05**	**0.0007**	**0.90**	**0.00**	**0.08**
ESRRB	**11.89**	**0.0006**	**0.0041**	**0.27**	**0.11**	**0.44**
NR2C1	**9.52**	**0.0020**	**0.0112**	**0.38**	**0.23**	**0.91**
ESRRA	**6.66**	**0.0098**	**0.0433**	**0.75**	**0.00**	**0.08**
PGR	5.67	0.0173	0.0633	0.71	0.05	0.18
NR5A2	4.59	0.0321	0.0946	0.30	0.21	0.00
NR4A3	4.47	0.0344	0.0946	0.19	0.39	0.10
THRA	4.02	0.0449	0.1098	0.025	1.91	0.00
LRT-4: M3(k = 2) *vs.* model D(k = 2)						
RORA	**47.02**	**7.04e-12**	**1.69e-10**	**0.11**	**0.26**	**3.40**
RARG	**15.88**	**6.76e-05**	**0.0008**	**0.90**	**0.00**	**0.08**
ESRRB	**11.89**	**0.0006**	**0.0045**	**0.28**	**0.11**	**0.44**
NR2C1	**9.59**	**0.0020**	**0.0117**	**0.81**	**0.03**	**0.28**
ESRRA	5.86	0.0155	0.0690	0.75	0.00	0.08
PGR	5.67	0.0173	0.0690	0.71	0.05	0.18
NR5A2	4.57	0.0325	0.0967	0.28	0.23	0.00
NR4A3	4.47	0.0344	0.0967	0.19	0.39	0.10
NR1H2	4.38	0.0363	0.0967	0.89	0.02	0.00
NR3C2	4.19	0.0407	0.0976	0.88	0.04	0.12

Genes having a q-value of <0.05 are shown in boldface type. The foreground (FG) branches are fully specified for each hypothesis in Figure 1. The null model for all LRTs assumes homogenous selection pressure for all branches ($\omega_{BG} = \omega_{FG}$). LRT-4 and LRT-5 have d.f. = 1. The q-value is the expected proportion of false discoveries expected if the single-test P-value is used as the boundary to control the FDR. p_{SHIFT} is the fraction of sites subject to a shift in selection pressure.

that statistical regularity conditions might not be relied upon to justify inference under the complex models employed in phase 2. For this reason, we added a novel use of the bootstrap to assess inference under those models in phase 2 (Bielawski *et al.* 2016).

Most NRs share a modular structure (Figure 3), and we investigated the extent to which evolutionary constraints were associated with those domain structures. We did not include the CTD, due to its absence from some NRs. In two cases (*NR0B1* and *NR0B2*) we were unable to assign sites to structural domains, so they were excluded from this analysis. We fit a fully heterogeneous codon model (denoted FE1 in Bao *et al.* 2007) to the remaining 46 genes in order to estimate the amount of among-domain variability in (i) the d_N/d_S ratio, ω; (ii) the transition-to-transversion rate ratio, κ; (iii) the branch length scale factor, c; and (iv) codon bias, π_j's. In this study we are primarily interested in ω as a measure of selection intensity. Aggregation of domain-specific parameter values indicated that NR domain evolution is relatively homogenous for both the transition-to-transversion rate ratio and the codon usage bias (Figure 4, A and B). In contrast, the intensity of selection (ω) differed substantially among domains; the DBD and the LBD were highly constrained by purifying selection and exhibited almost no nonsynonymous changes (Figure 4C). The DBD was the most conserved, with median ω of just 0.004 (Q1 = 0.001; Q3 = 0.009), and the LBD had median ω of 0.014 (Q1 = 0.007; Q3 = 0.041). Although still dominated by purifying selection, the NTD (median $\omega = 0.11$) and the HD (median $\omega = 0.07$) tended to exhibit more nonsynonymous changes. Moreover, for some genes, these two domains exhibited considerable divergence among primates (NTD Q3: $\omega = 0.25$; HD Q3: $\omega = 0.20$). Interestingly, we also observed signs of among-domain variability in scale parameters (c's), suggesting the possibility of different synonymous rates among domains, or a history of recombination (Figure 4D). Application of LRTs to directly assess model complexity for each gene (Bao *et al.* 2007) yielded similar results (File S1). In response to the observed variation in c's, we added two more robustness analyses to phase 2. Specifically, we applied (i) a genetic algorithm to screen alignments for evidence of recombination (Kosakovsky Pond *et al.* 2006) and (ii) a multilayer (DNA/RNA and protein) codon model to assess among-site variability in the baseline rate of DNA/RNA substitution (Rubinstein *et al.* 2011).

Phase 2: Survey for hominid-specific alterations in the intensity of natural selection

Survey design: The objectives of this phase of the research were to (i) identify NRs that could have played a role in hominid evolution and (ii) choose the best single candidate for AGR-based laboratory investigation. We designed a survey to detect signal for sites having either an episodic change in selection pressure (Figure 1, H_1-H_3), or a long-term shift in selection pressure (Figure 1, H_4 and H_5), associated with the evolution of hominids. The criterion for an episodic change was a significant result under LRT-1 or LRT-2, and the criterion for a long-term shift was a significant result under LRT-3

Table 3 Robustness of inferences about long-term shifts in the intensity of selection pressure at the origin of the great apes (H_4)

		Model C				Model D			
Gene	Analysis	p_{SHIFT}	ω_{BG}	ω_{FG}	LRT	p_{SHIFT}	ω_{BG}	ω_{FG}	LRT
RORA	Original	0.11	0.26	3.40	P = 7.0e-12	0.11	0.26	3.40	P = 7.0e-12
	MG94	0.11	0.30	3.90	P = 8.0e-12	0.11	0.30	3.90	P = 8.0e-12
	Gene tree	0.12	0.24	3.03	P = 1.9e-11	0.12	0.24	3.03	P = 1.9e-11
	Bootstrap	[0.04–0.17]	[0.04–0.42]	[0.83–9.53]	N.A.	[0.06–0.17]	[0.11–0.43]	[1.44–8.32]	N.A.
RARG	Original	0.90	0	0.08	P = 6.8e-05	0.90	0	0.08	P = 6.8e-05
	MG94	0.89	0	0.10	P = 7.1e-05	0.89	0	0.10	P = 7.1e-05
	Gene tree	0.90	0	0.08	P = 6.8e-05	0.90	0	0.08	P = 6.8e-05
	Bootstrap	[0.02–0.96][a]	[0–6e-3]	[0.02–3.06]	N.A.	[0.77–0.99]	[0–8e-3]	[0.03–0.15]	N.A.
ESRRB	Original	0.27	0.11	0.44	P = 0.0006	0.27	0.11	0.44	P = 0.0006
	MG94	0.27	0.14	0.59	P = 0.0007	0.27	0.15	0.59	P = 0.0007
	Gene tree	0.27	0.11	0.44	P = 0.0006	0.27	0.11	0.44	P = 0.0006
	Bootstrap	[0.14–0.75]	[0–0.16]	[0.07–1.00]	N.A.	[0.14–0.45]	[0.04–0.17]	[0.19–1.07]	N.A.
NR2C1	Original	0.38	0.23	0.91	P = 0.002	0.81	0.03	0.28	P = 0.002
	MG94	0.37	0.25	1.02	P = 0.001	0.36	0.27	1.0	P = 0.002
	Gene tree	Match	Match	Match	N.A.	Match	Match	Match	N.A.
	Bootstrap	[0.05–0.71]	[5e-3–0.36]	[0.21–2.86]	N.A.	[0.65–0.94]	[0–0.05]	[0–0.70]	N.A.
ESRRA	Original	0.75	0	0.08	P = 0.010	0.75	0	0.08	P = 0.015
	MG94	0.69	0	0.09	P = 0.010	0.69	0	0.09	P = 0.014
	Gene tree	0.75	0	0.08	P = 0.008	0.75	0	0.08	P = 0.013
	Bootstrap	[0.24–1]	[0–0.03]	[0–0.21]	N.A.	[0.162–1][a]	[0–0.02]	[0–0.35]	N.A.

Robustness and bootstrapping analyses were carried out for genes having at least one q-value of <0.05. For analyses under the gene tree, a match designation indicates that the gene tree was identical to the organismal tree.

[a] Bimodal distribution.

or LRT-4. Because we tested each hypothesis in 48 different alignments, the single test significance level will not provide adequate control over the probability of making one or more type I errors. We used the FDR criterion of Storey (2002) to identify the subgroup of genes related by a given evolutionary scenario (Figure 1, H_1–H_5). The alignments for every gene within a subgroup were visually inspected and verified a second time by a different coauthor. The subgroups were then screened according to a suite of reliability and robustness analyses.

Assessing the signal for linage specific episodes of adaptive evolution: We formally tested each NR for sites having an episode of altered selection pressure at the origin of the great apes (H_1), the origin of the human–chimpanzee clade (H_2), or along the human lineage (H_3) (Figure 1). LRT-1 and LRT-2 revealed little evidence of episodic evolution (Table 1). One gene (*NR0B2*) was significant for a change in selection pressure at the origin of the great apes; no genes were significant at the origin of human–chimpanzee clade; and four genes (*NR1D1*, *PPARG*, *NR2C1*, and *PGR*) were significant for a change along the human lineage. However, only one (*NR1D1*) was significant after controlling for false discoveries (using a q-value of <0.05 as the selection criterion). We used nonparametric bootstrapping to quantify uncertainty and check for instabilities in the MLEs of parameter values for *NR1D1*. The estimated distribution for the fraction of sites under episodic evolution (p_{FG}) was strongly bimodal, with considerable density at zero (Figure 5), indicating a clear departure from the expected limiting properties of this MLE. This instability indicates that the conditions necessary for reliable MLEs and LRTs (see Self and Liang 1987) cannot be assumed for *NR1D1*. Thus, we found no reliable evidence for episodic adaptive evolution in primate NRs.

Assessing the signal for long-term shifts in the intensity of selection pressure in the great ape clade (H_4): In contrast to our analyses of episodic evolution, clade-site models (Bielawski and Yang 2004) uncovered evidence for a nontrivial subgroup of genes having a shift in selection pressure in some lineages. We employed both LRT-3 and LRT-4 to test for a shift at the origin of the great apes (Figure 1, H_4). Nine genes were significant for LRT-3, with five remaining significant after controlling the FDR (Table 2: q-value <0.05). Ten genes were significant for LRT-4, but only four had a q-value of <0.05 (Table 2). The same four genes (*RORA*, *RARG*, *ESRRB*, and *NR2C1*) also had a q-value of <0.05 for LRT-3, suggesting strong signal for a long-term shift within this subgroup. The fraction of sites subject to a shift in selection pressure (p_{SHIFT} in Table 2) ranged from ~11% (*RORA*) to 90% (*RARG*) under model D. The estimate of p_{SHIFT} for *RARG* is extreme, suggesting the possibility of a large estimation error, or even a failure to meet statistical regularity conditions.

Further interpretation of the MLEs should always be treated with caution without additional assessment. Therefore, we carried out a series of robustness analyses for the five genes that had a q-value of <0.05 under H_4 (*RORA*, *RARG*, *ESRRB*, *NR2C1*, and *ESRRA*). None of these genes exhibited evidence of recombination, as inferred under the GARD-MBP method (File S5). All five genes had highly consistent MLEs and LRTs under the MG94-style codon models, as well as under gene

trees that differed from the organismal tree (Table 3). Note that gene trees differed from the assumed organismal tree for all genes except *NR2C1*.

Next we used bootstrapping to assess the MLE distribution of the five genes. The MLE distributions for three genes (*RORA*, *ESRRB*, and *NR2C1*) were unimodal and bell-shaped (allowing for boundaries such as the prohibition of negative frequencies). Although there is uncertainty associated with the point estimates for the model parameters (Table 3), the bootstrap distributions upheld the signal for sites subject to divergent selection pressure in all three genes ($p_{SHIFT} > 0$), as well as the signal for positive selection within *RORA* (Table 3). In contrast, bootstrapping revealed that the MLE distribution for p_{SHIFT} was bimodal for *RARG* and *ESRRA*. This indicates a clear departure from the asymptotic distributional properties expected when regularity conditions are satisfied for this MLE. We used multiple analyses started from different initial parameter values to confirm that these bimodal distributions were not due to suboptimal peaks in likelihood. This indicates that conditions necessary for reliable MLEs and LRTs cannot be assumed for *RARG* and *ESRRA*, despite having *q*-values of <0.05. We consider the signal for a shift in selection pressure at the origin of the great apes to be reliable for *RORA*, *ESRRB*, and *NR2C1* only.

Assessing the signal for long-term shifts in the intensity of selection pressure in the human–chimpanzee clade (H_5): We also employed LRT-3 and LRT-4 to test for a shift in selection pressure at the origin of the human–chimpanzee clade (Figure 1, H_5). As displayed in Table 4, eight genes were significant for LRT-3 and LRT-4. After controlling for false discovery (*q*-value <0.05), five genes remained under LRT-3 (*NR2C1*, *NR1D1*, *PGR*, *NR2E3*, and *ESRRB*) and three genes remained under LRT-4 (*NR2C1*, *NR2E3*, and *PGR*), suggesting a strong signal for a long-term shift within the latter subgroup. Interestingly, estimates of the ω distribution under clade-site models indicate a substantial fraction of sites (22–36%) evolving under positive selection since the origin of the human–chimpanzee clade.

We performed additional assessments of robustness and bootstrapped the MLEs for the five candidate genes. Two genes (*NR1DI* and *ESRRB*) were excluded because bootstrapping revealed bimodal MLE distributions suggestive of parameter estimate instabilities (Table 5). *NR2E3* was excluded due to significant signal for recombination obtained by using the GARD method (File S5). Inference of selection was generally robust to assumptions about among-site variation in the baseline DNA/RNA substitution rate (File S6). Although *PPARG* did not have a *q*-value of <0.05, it is noteworthy that ω's were very sensitive to assumptions about the baseline DNA/RNA substitution rate in this gene (File S6). Such sensitivity highlights the importance of assessing this in all candidate genes.

The remaining two candidates (*NR2C1* and *PGR*) had unimodal and approximately bell-shaped MLE distributions. Parameter estimates suggesting positive selection in the human–chimpanzee clade were consistent across alternative codon models (models C and D), modeling frameworks (GY94 *vs.* MG94), and tree topologies (gene *vs.* organism) (Table 5). Bootstrapping corroborated the signal for a fraction of sites subject to divergent selection pressure among the human–chimpanzee clade, as the estimates of p_{SHIFT} were always > 0. The bootstrap also indicated substantial density for $\omega_{FG} > 1$; however, it was not 100% of the distribution (models C/D: 90/94% $\omega_{FG} > 1$ for *NR2C1*; 95/98% $\omega_{FG} > 1$ for *PGR*). Because any change in the distribution of ω, even when $\omega < 1$, is an indicator of functional divergence at the molecular level (Forsberg and Christiansen 2003; Bielawski and Yang 2004) we conclude that the signal for functional divergence is strong in *NR2C1* and *PGR* under H_5.

Identifying the top candidates for AGR-based experimental assays and selection of NR2C1 for further investigation: When the H_4 and H_5 results are combined, we are left with a set of four genes (*ESRRB*, *NR2C1*, *PGR*, and *RORA*) whose evolution is statistically associated with the origin of the great ape and human–chimpanzee clades. AGR-based characterization of each gene is not within the scope of this study, nor is each one an equally promising candidate. Hence, we carried out a subjective ranking and chose *NR2C1* as our top candidate for further AGR-based experimental assays. Our ranking was based on our assessment of the signal for a change in selection pressure, the extent to which the biological role of each gene has been characterized, and the capacity to carry out *in vitro* assays for functional effects of amino acid substitutions. The three genes that we did not

Table 4 Significant LRTs for a subset of sites having experienced a shift in the intensity of selection pressure at the origin of the human–chimpanzee clade (H_5)

Gene	$2\delta l$	*p*	*q*	p_{SHIFT}	ω_{BG}	ω_{FG}
LRT-3: M2a-rel *vs.* model C						
NR2C1	**12.57**	**0.0004**	**0.0126**	**0.36**	**0.25**	**2.06**
NR1D1	**8.70**	**0.0032**	**0.0384**	**0.002**	**0.00**	**126.89**
PGR	**8.47**	**0.0036**	**0.0384**	**0.30**	**0.54**	**2.57**
NR2E3	**7.90**	**0.0049**	**0.0395**	**0.22**	**0.41**	**3.39**
ESRRB	**7.23**	**0.0071**	**0.0458**	**0.27**	**0.12**	**0.61**
PPARG	5.01	0.0251	0.1341	0.12	0.43	1.86
ESR2	4.07	0.0437	0.1820	0.74	0.04	0.40
NR3C2	4.00	0.0455	0.1820	0.12	0.52	2.71
LRT-4: M3(k = 2) *vs.* model D(k = 2)						
NR2C1	**12.55**	**0.0004**	**0.0111**	**0.35**	**0.26**	**2.13**
NR2E3	**9.38**	**0.0022**	**0.0306**	**0.22**	**0.41**	**3.39**
PGR	**8.47**	**0.0036**	**0.0336**	**0.30**	**0.54**	**2.57**
ESRRB	7.23	0.0071	0.0500	0.27	0.12	0.61
PPARG	5.01	0.0251	0.1408	0.12	0.43	1.86
ESR2	4.12	0.0424	0.1629	0.79	0.05	0.40
NR3C2	4.03	0.0448	0.1629	0.11	0.54	2.72
NR1D1	3.96	0.0465	0.1629	0.06	0.81	4.88

Genes having a *q*-value of <0.05 are shown in boldface type. The foreground (FG) branches are fully specified for each hypothesis in Figure 1. The null model for all LRTs assumes homogenous selection pressure for all branches ($\omega_{BG} = \omega_{FG}$). LRT-4 and LRT-5 have d.f. = 1. The *q*-value is the expected proportion of false discoveries expected if the single-test *P*-value is used as the boundary to control the FDR. p_{SHIFT} is the fraction of sites subject to a shift in selection pressure.

Table 5 Robustness of inferences about long-term shifts in the intensity of selection pressure at the origin of the human–chimpanzee clade (H_5)

		Model C				Model D			
Gene	Analysis	p_{SHIFT}	ω_{BG}	ω_{FG}	LRT	p_{SHIFT}	ω_{BG}	ω_{FG}	LRT
NR2C1	Original	0.36	0.25	2.06	$P = 0.0004$	0.35	0.26	2.13	$P = 0.0004$
	MG94	0.37	0.25	1.02	$P = 0.0015$	0.36	0.27	1.04	$P = 0.0017$
	Gene tree	Match	Match	Match	N.A.	Match	Match	Match	N.A.
	Bootstrap	[0.17–0.57]	[0.11–0.42]	[0.31–8.32]	N.A.	[0.18–0.47]	[0.18–0.41]	[0.4–9.41]	N.A.
NR1D1	Original	0.002	0	126.89	$P = 0.0032$	0.06	0.81	4.88	$P = 0.0465$
	MG94	0.002	0	162.68	$P = 0.0057$	0.07	0.74	4.19	$P = 0.0511$
	Gene tree	Match	Match	Match	N.A.	Match	Match	Match	N.A.
	Bootstrap	[0–8e-3]	[0–2.12][a]	[8.84–99]	N.A.	[0.02–0.13]	[0.38–1.43]	[0–22.31]	N.A.
PGR	Original	0.30	0.54	2.57	$P = 0.0036$	0.30	0.54	2.57	$P = 0.0036$
	MG94	0.30	0.59	2.70	$P = 0.0049$	0.30	0.60	2.68	$P = 0.0049$
	Gene tree	0.31	0.54	2.54	$P = 0.0011$	0.31	0.54	2.54	$P = 0.0038$
	Bootstrap	[0.18–0.5]	[0.22–0.7]	[0.65–5.11]	N.A.	[0.22–0.42]	[0.41–0.68]	[0.77–8.37]	N.A.
NR2E3	Original	0.22	0.41	3.39	$P = 0.0049$	0.22	0.41	3.39	$P = 0.0022$
	MG94	0.21	0.45	3.97	$P = 0.0051$	0.21	0.45	3.94	$P = 0.0015$
	Gene tree	***	***	***	N.A.	***	***	***	N.A.
	Bootstrap	[0.13–0.36]	[0.26–0.71]	[0.36–99]	N.A.	[0.11–0.34]	[0.27–0.69]	[0.6–13.98]	N.A.
ESRRB	Original	0.27	0.12	0.61	$P = 0.0071$	0.27	0.12	0.61	$P = 0.0071$
	MG94	0.26	0.17	0.80	$P = 0.0081$	0.26	0.17	0.80	$P = 0.0081$
	Gene tree	0.27	0.12	0.61	$P = 0.0076$	0.27	0.13	0.61	$P = 0.0076$
	Bootstrap	[0.15–0.84][a]	[0–0.21][a]	[0.07–1.92]	N.A.	[0.1–0.40]	[0.06–0.22]	[0.07–1.84]	N.A.

Robustness and bootstrapping analyses were carried out for genes having at least one q-value of <0.05. For analyses under the gene tree, a match designation indicates that the gene tree was identical to the organismal tree. Analyses that were impossible because the gene tree topology prevented specification of H5 are indicated by ***.
[a] Bimodal distribution.

select for further investigation (*ESRRB*, *PGR*, and *RORA*) remain, nonetheless, very strong candidates for future work. Further details about our ranking of those genes are provided in File S7.

Our top candidate, *NR2C1*, belongs to a subtype of NRs known as orphan receptors for which the endogenous ligand (if any) has yet to be identified (Lee and Chang 1995). Originally named testicular receptor 2 (*TR2*) because it was first isolated from human testis and prostate (Chang and Kokontis 1988; Anderson *et al.* 2012), its expression in ES cells and in pluripotent cell culture lines indicates it plays a role in early embryonic development (Hu *et al.* 2002). It is one of a handful of genes implicated in the regulation of the pluripotentiality of stem cell populations in the embryo and in neural stem cells in particular (Lee and Chang 1995; Hu *et al.* 2002; Lee *et al.* 2002; Shyr *et al.* 2009). In addition, *NR2C1* has been shown to regulate the expression of *Oct4* and *Nanog*, two transcription factors essential for maintaining the pluripotentiality of embryonic stem cells (Pikarsky *et al.* 1994; Niwa *et al.* 2000; Boiani 2002). Examination of the level of *NR2C1* expression in induced pluripotent stem cells (iPSCs) generated from human melanocytes, fibroblasts, and hepatocytes from previously published experiments (Gene Expression Omnibus (GEO) accession no. GDS3867; Ohi *et al.* 2011) similarly shows that *NR2C1* expression is increased in each iPSC line relative to its nonpluripotent parental cell type. Furthermore, expression of *NR2C1* in chimpanzee iPSCs is increased relative to their undifferentiated parental cell lines (Gallego Romero *et al.* 2015). Given its role in maintaining the pluripotentiality of stem cells and in neural differentiation, and a signal for positive selection within the human–chimpanzee clade (H_5), we hypothesized that *NR2C1* has played a role in the evolution of a developmental system(s) relevant to the anatomical and physiological characteristics that distinguish humans and chimpanzees from the other great apes. In the next phase of this study, we take the first steps to evaluate this hypothesis by investigating if amino acid substitutions within *NR2C1* since the LCA of humans and chimpanzees have affected its capacity to regulate pluripotentiality.

Phase 3: Inference of an ancestral NR2C1 sequence, synthesis of an ancestral protein, and in vitro assessment of its gene regulatory effects

Experimental design: In this phase of the investigation, we set out to determine whether the *NR2C1* amino acid substitutions unique to either modern humans (*hNR2C1*), chimpanzees (*cNR2C1*), or the inferred LCA of humans and chimpanzees–bonobos (*aNR2C1*) alter the ability of the gene to maintain pluripotentiality. More specifically, by heterologously expressing these proteins in embryonic stem cells, we tested whether these sequences differ (i) in their ability to maintain pluripotentiality, (ii) in their relative ability to regulate transcripts associated with pluripotentiality (*i.e.*, *Oct4* and *Nanog*), and (iii) whether they differentially regulate the transcription of a promoter element associated with NR-mediated signaling (Lucas *et al.* 1991) and differentiation (Zimmer and Magnuson 1990). Any change in the efficiency of *NR2C1* as a transcriptional activator of pluripotentiality or of differentiation-related genes, or any change in its ability

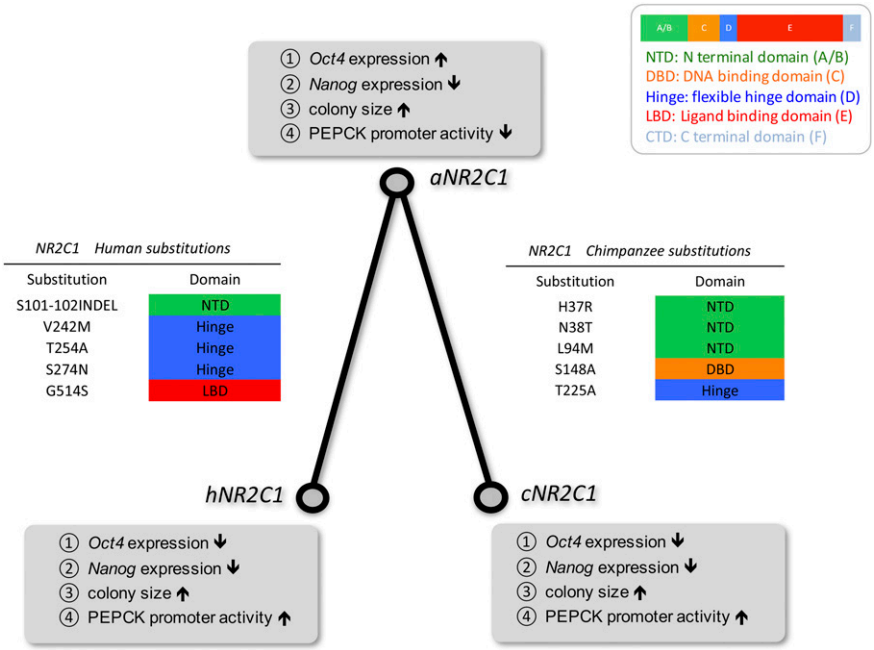

Figure 6 Amino acid substitutions along the human and chimpanzee lineages and their impact on four different measures of molecular phenotype. Lineage-specific amino acid substitutions were inferred by comparing the ancestral state with the highest posterior probability to the extant state at a site in human and chimpanzee *NR2C1*. The location of each substitution is given with respect to alignment position and structural domain. The cDNA corresponding to the ancestral (*aNR2C1*) and extant (*hNR2C1* and *cNR2C1*) gene sequences was synthesized, and experimental assays were employed to measure four different aspects of molecular phenotype (*Oct4* expression, *Nanog* expression, colony size, and *Pepck* promoter activity). The results of those assays are summarized within the gray boxes at the three nodes of the two-taxon phylogeny. The up and down arrows indicate the relative effect of the amino acid substitutions on the molecular phenotype between the ancestral and extant variants.

to maintain pluripotentiality in a stem cell pool, could be evolutionarily significant. Such change, even if relatively small, could underlie substantial changes in overall anatomical and physiological characteristics.

We used maximum likelihood estimates of model parameters to infer the LCA of humans and chimpanzees–bonobos (*aNR2C1*) according to empirical Bayes posterior probabilities. The ancestral amino acid sequence, along with additional details about the methods of ancestral state reconstruction, are provided in File S3. We then sequenced *NR2C1* transcripts from multiple primates and used these in combination with available sequence data to exclude the possibility that some of the inferred amino acid substitutions might represent polymorphisms or errors in the prediction of the messenger RNA (mRNA) sequences (File S4). Figure 6 gives the amino acid substitutions along the human and chimpanzee lineages implied by the inferred ancestral sequence. The cDNA for the ancestral (*aNR2C1*) and extant (*hNR2C1* and *cNR2C1*) genes was synthesized, and all subsequent assays reflect the amino acid states corresponding to those sequences.

Evidence for functional divergence in the transcriptional regulation of pluripotentiality genes by NR2C1 gene variants: Pluripotency is the transient attribute of single embryonic cells to generate all cell lineages of the developing and adult organism. During embryonic development, pluripotent stem cell populations form organs and tissues by differentiating in a stepwise fashion. One way of assessing this process is by following the expression of molecular markers of differentiation, such as *Oct4* and *Nanog*. *Oct4* and *Nanog* are important regulators of pluripotency and cell lineage commitment and are essential for establishing and maintaining pluripotentiality (Nichols *et al.* 1998; Mitsui *et al.* 2003). As mouse *NR2C1* (*mNR2C1*) has been shown to regulate the expression of *Oct4* and *Nanog* in mouse ES cells (Pikarsky *et al.* 1994; Niwa *et al.* 2000; Boiani 2002), we investigated whether the amino acid divergence between the *NR2C1*s of human, chimpanzee, and the inferred LCA might alter its ability to regulate these key pluripotentiality genes.

We expressed *hNR2C1*, *cNR2C1*, and *aNR2C1* in mouse ES cells under the control of a strong promoter along with a siRNA construct that knocks down endogenous expression of *mNR2C1*. Cotransfected cells were selected by addition of antibiotics (see *Materials and Methods*). The results of the knockdown/overexpression experiments were compared to both a negative knockdown control (which received the siRNA construct and an empty heterologous expression vector) and a nonsense positive control [which received a control (nonsense) siRNA and an empty heterologous expression vector]. Our goal was to knock down *NR2C1* and then replace it with the same gene from different species. Mouse ES cell lines are very well-defined pluripotent lines, and the three tested variants are all effectively the same evolutionary distance from mouse. Using the mouse line also allowed us to knock down the endogenous *NR2C1* expression using siRNAs, and then reexpress the test *NR2C1* constructs without any interference from the siRNAs. The human codon optimization of the synthetically generated chimpanzee and ancestral *NR2C1* constructs was performed to ensure that all constructs were expressed with equivalent efficiency and equally divergent from the endogenous mouse sequence that was being silenced from the mouse knockdown constructs.

The results of our analysis demonstrate a 75% decrease in *NR2C1* expression in the knockdown cohort, thus confirming *NR2C1* was knocked down by our siRNA construct (Figure 7A). We also confirmed an earlier report (Shyr *et al.* 2009) that *Oct4* and *Nanog* had significantly reduced expression in

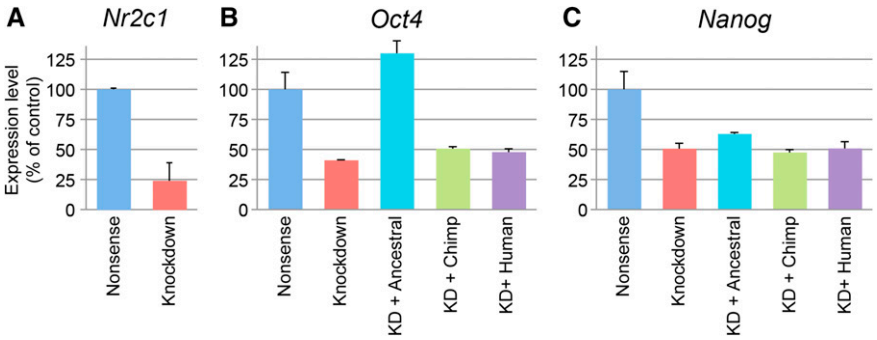

Figure 7 Effect of *NR2C1* knockdown on pluripotentiality-related gene expression in ES cells. (A) ES cells transfected with a m*NR2C1* knockdown vector show significantly reduced *NR2C1* expression relative to ES cells receiving a nonsense control vector (24%, $P < 0.01$). (B) Cotransfection of a*NR2C1* with knockdown vector rescues expression of *Oct4* (130%), but *hNR2C1* and *cNR2C1* do not. (C) Normal expression of Nanog is not rescued by cotransfection with any of the *NR2C1* plasmids.

the knockdown cohort (41 and 51% reductions relative to the positive control). We then evaluated whether overexpression of the *NR2C1* variants could rescue the reduced expression of *Oct4* and *Nanog*. We found that the ancestral gene variant rescued *Oct4*, even reaching levels above that observed in control-transfected cells, with a 130% increase in expression compared to the negative control. We did not observe any rescue by the chimpanzee or human variants (Figure 7B), as the expression levels of *Oct4* remained at the levels observed in the knockdown cohort. None of the *NR2C1* variants rescued *Nanog* (Figure 7C).

These results suggest two key points. First, the lack of rescue by any primate variant suggests functional divergence in either the transcription factor or in the mouse *Nanog* promoter. Such divergence is plausible, given that primates and rodents have been evolving independently for ~75 million years; however, it is by no means inevitable (Wasserman *et al.* 2000; Liu *et al.* 2004). Second, the fact that only the ancestral variant rescues *Oct4* expression indicates that the ancestral and extant forms of *NR2C1* differ in their ability to bind to, and/or transcriptionally activate the mouse *Oct4* promoter. Our results raise the possibility that the inferred ancestral form of *NR2C1* may have an activity level closer in function to the mouse variant than to its human and chimpanzee counterparts. Thus, our findings imply that considerable functional divergence in *NR2C1* occurred sometime within the last 8–13 million years of hominid evolution and that amino acid substitutions at no more than five sites was sufficient to alter the transcriptional properties of this key pluripotentiality gene. Additional work is required to determine whether the divergence of *NR2C1* was a unique evolutionary event or if the regulatory capacity of *NR2C1* is relatively plastic and has been more widely modified within primates.

In vitro assays of pluripotentiality suggest differential activity of hNR2C1, cNR2C1, and aNR2C1: Given that we demonstrated functional divergence between the ancestral and extant forms of primate *NR2C1* to regulate *Oct4*, we predicted that the knockdown of endogenous *mNR2C1* and its replacement with its primate variants could result in a detectable change in the pluripotentiality of transfected ES cells. As ES cells must remain pluripotent in order to undergo self-renewal under their standard culture conditions, the ability of ES cells to form new clonal colonies is a reliable proxy of their pluripotentiality (Chambers and Smith 2004). We exploited this property of ES cells to assess and compare the phenotype of pluripotentiality in the extant and ancestral forms of the three primate *NR2C1* gene variants.

We selected 12 clonal ES cell cultures for each of the above plasmid combinations and plated these cells at a limiting dilution to assess their ability to generate new pluripotent clones. We assayed the number of clones formed for each replicate and then assessed the average size of the clonal colonies by harvesting, dissociating, and counting 10 random clones for each well. As it is likely that factors influencing the pluripotentiality of the ES cell population will act gradually over time by limiting the number of cell divisions over which the cells can successfully self-renew, we repeated this assay by replating the dissociated cells across four iterations. Of the initial 60 replicates, 58 survived through the duration of the experiment (two did not survive the initial passage from a single selected clone and are not included in the analysis). A representative photo of colonies is provided in Figure S1.

Although the knockdown of *NR2C1* quantitatively and substantially reduces the expression of *Oct4* and *Nanog*, by assessing the number of clones formed, we found that the lack of *NR2C1* does not impair their self-renewal ability in our *in vitro* assays; indeed, even at the end of the experiment, numerous ES cell colonies were present in knockdown cultures, and both knockdown and negative control colonies were morphologically normal. Similarly, overexpression of the human, chimpanzee, and inferred ancestral variants of *NR2C1* does not significantly alter their ability to self-renew. Thus, it appears that neither the loss of endogenous *NR2C1* nor the ectopic expression of the primate variants leads to outright differentiation of ES cells as is observed following a direct knockdown of *Oct4* or *Nanog* (Hay *et al.* 2004; Zaehres *et al.* 2005) and leads us to conclude that the role of rodent and primate *NR2C1* in regulating pluripotentiality appears to be more complex than simply acting as an upstream regulator of *Oct4/Nanog*.

Interestingly, a difference was observed in the average size of the clonal colonies, which suggests that *NR2C1* may regulate other aspects of differentiation and proliferation. Knockdown colonies were approximately one-third smaller than the positive control colonies (68%; $P = 0.04$ by Student's *t*-test). Expression of any of the chimpanzee, human, or ancestral variants of *NR2C1* reversed this effect (114, 120, and

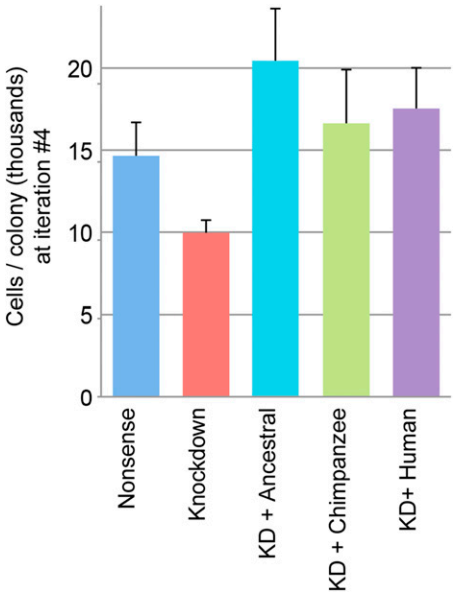

Figure 8 The loss of *NR2C1* activity appears to impact the ES cells proliferative ability; knockdown colonies appear approximately one-third smaller than the positive control colonies (68%; $P = 0.04$ by Student's *t*-test). Expression of any of the chimpanzee, human, or ancestral variants of *NR2C1* reverses this effect (114, 120, and 140%, respectively).

140%, respectively, although differences among the variants are not significant by one-way ANOVA). As stem cells differentiate, their cell cycle times change (*e.g.*, as neural stem cells progress through the stages of neural differentiation, the length of their cell-cycle increases). Therefore, for a given stem cell population, the rate of proliferation is effectively defined by their level of differentiation. Collectively, our results raise the possibility that evolutionary modulation of *NR2C1* may be related to the fundamental properties of the stem cell population. It is interesting that the ancestral variant, which rescues *Oct4* expression, exhibited the largest colony size in our assays (Figure 8).

In vitro assays suggest hNR2C1, cNR2C1, and aNR2C1 differentially regulate a differentiation-associated promoter: The primary function of *NR2C1* is to act as a transcription factor by binding to specific sites in the promoters of target genes and then either activating or repressing their expression. Previous studies identified one such candidate target of *NR2C1* by showing the regulation of the phosphoenolpyruvate carboxykinase (*Pepck/Pck1*) promoter using a luciferase reporter assay (Shyr *et al.* 2009). Expression of the *Pepck* gene increases in numerous cell types as development proceeds, including in the developing neural tube (Zimmer and Magnuson 1990) and liver endothelial cells (Gruppuso *et al.* 1999). As *Pepck* is a key metabolic enzyme, changes in its expression likely reflect changes in metabolic activity during differentiation. Such changes in metabolic activity are evident even in the earliest stages of stem cell differentiation, and may in fact be intimately linked to pluripotentiality (Mandal *et al.* 2011).

To examine whether this function differs among the three gene variants, we assayed the ability of the *NR2C1* variants to modulate transcription using an *in vitro* luciferase reporter gene. We transfected HEK-293 cells with the plasmids encoding *aNR2C1*, *cNR2C1*, and *hNR2C1* as well as the *Pepck*-LUC reporter plasmid. After an overnight culture we lysed the cells to measure luciferase activity. Higher light output in this assay is indicative of increased promoter activity. HEK-293 is a transformed human cell line that normally expresses modest levels of *Pepck*, as evidenced in this assay by significant basal activity of the *Pepck* promoter. We found that *hNR2C1* shows virtually no ability to modulate this activity (Figure 9). In contrast, *aNR2C1* showed a significant and substantial repression of the basal activity of the *Pepck* promoter (64.7% relative to control, $P < 0.0001$ by ANOVA with Dunnett's post-test). This repression was unique to the ancestral variant, as the *cNR2C1* transfection caused a significant increase (112% relative to control, $P = 0.03$ by ANOVA with Dunnett's post-test). Thus, it appears that the three primate variants of *NR2C1* have differential abilities to regulate transcription from this promoter element.

As there are relatively few amino acid changes between these variants, it is possible to speculate on which changes may underlie the differential transcriptional response. The human and the ancestral variants have identical DBDs, but different transcriptional activities. As the HD is thought to have minimal effect on NR function, substitutions at the three sites identified in that domain (Figure 6) are not likely to have a direct effect on transcriptional activity. In comparing the ancestral and chimpanzee variants, the candidate substitutions lie at sites in either the NTD or DBD. Because both ancestral and chimpanzee *NR2C1* have transcriptional activity, it seems unlikely that the difference in activity is due to an inability of the DBD of one variant to bind to the *Pepck* promoter. Thus, it is likely that the three substitutions in the NTD are the strongest candidates for direct modulation of the transcriptional activity along the chimpanzee lineage.

Conclusions

We found that rigorous screening of candidate genes for (i) robustness to model assumptions and (ii) reliability of the MLEs are critical components of experimental design in an evolutionary survey. While it is common practice to employ alternative formulations of codon models to assess robustness of signal for positive selection (*e.g.*, consistency between model A and model B), further assessment of the assumptions upon which such models are built is rarely pursued. We expect that robustness analyses will have a greater impact on studies of lineages more divergent than primates. More critical to future investigations of primate molecular evolution was the finding of instabilities in the MLE distributions for several genes [*ESRRA*, *ESRRB* (under H_5), *NR1D1*, and *RARG*], which indicated that the conditions for reliable interpretation of the MLEs and LRTs could not be assumed. We suggest that consistency among alternative codon models, although desirable, should not be taken as a "talisman of

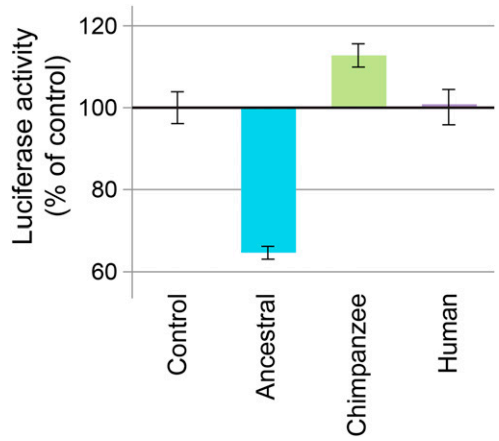

Figure 9 Reporter assay used to assess regulation of a differentiation-associated promoter (*Pepck*-luciferase). While the human *NR2C1* (*hNR2C1*) construct did not exhibit any ability to modulate the activity of the *Pepck* promoter, the ancestral *NR2C1* (*aNR2C1*) displayed a significant repression of the basal activity of the promoter, and the chimpanzee *NR2C1* (*cNR2C1*) transfection caused a significant increase in the activity of the *Pepck* promoter. Luciferase activity measurements are the average of eight assays and normalized to nonsense control assay values.

reliability" because it provides no indication of whether the underlying requirements for inference have been met.

From an initial set of 48 NRs, we identified a set of four genes [*ESRRB* (under H_4), *PGR*, *RORA*, and *NR2C1*] whose evolution was statistically associated with the origin of the great ape or the human–chimpanzee clades and that we consider good candidates for further investigation via ancestral state reconstruction, gene synthesis, and laboratory analyses. These results, which were derived from clade-site models C and D, contrast with those derived from branch-site models A and B, where we were unable to identify any candidates for further investigation. This difference is, in part, due to the different evolutionary assumptions made by the models. Models C and D assume that some sites experience a shift in selection pressure independent of whatever selection pressure happens to be acting on other sites. Models A and B, on the other hand, assume that some sites experience an episodic shift away from an ancestral level that is equivalent to the level at other sites. Thus, our evolutionary survey supports the notion that signal relevant to functional divergence can go undetected when only a single family of models is used to characterize their evolution (Schott *et al.* 2014).

Our experimental characterization of human, chimpanzee, and ancestral *NR2C1*'s transcriptional activities validated the evolutionary signal for functional divergence derived from the clade-site models. Comparison of *Oct4* expression in the three gene variants that had *NR2C1* knocked down demonstrate the ancestral variant has a 130% increase in expression of *Oct4*. This observation is consistent with divergence in the regulation of *Oct4*. As none of the primate variants rescued *Nanog*, our assays suggest additional divergence must have occurred between primates and rodents in either the promoter of the mouse *Nanog* gene or within the gene itself. Differences among the three *NR2C1* variants also modulated the promoter activity for *Pepck*, an enzyme whose metabolic activities are closely linked to development and differentiation (Zimmer and Magnuson 1990; Gruppuso *et al.* 1999; Mandal *et al.* 2011). Our results indicate that differential abilities to bind and regulate transcription from this promoter element are a result of substitutions at a relatively small subset of sites, likely in the NTD. Lastly, our assays suggest that reduced *NR2C1* levels in the knockdown construct are associated with reduced proliferation, suggesting that its activity positively regulates ES cell proliferation. The fact that the average colony size of the *aNR2C1*-transfected cells was the highest provides an intriguing hint that *aNR2C1* may have a more potent ability to rescue this proliferation defect than *hNR2C1* and *cNR2C1*.

Our results are consistent with the hypothesis that evolution of *NR2C1* is associated with lineage-specific differences in several aspects of stem cell populations related to the developmental phenotype of pluripotentiality. Given that neurogenesis occurs over a very limited period of development, evolutionary modulation of pluripotentiality is likely to have more profound effects on the proliferation of central nervous system neurons than on other organs or tissues. Thus, we propose that evolution of *NR2C1* could be associated with anatomical or physiological characteristics that distinguish humans from other great apes. Finally, analysis of *NR2C1* expression in the mouse embryo illustrates that it is robustly expressed in placodally derived neuroectodermal stem cells and presumptive neural-crest-derived mesenchymal stem cells (Baker *et al.* 2016). Further investigation as to whether this gene can regulate self-renewal and/or pluripotentiality in populations of stem cells other than ES cells such as in cultures of neural stem cells, or in iPSCs generated from non-human primates (Marchetto *et al.* 2013; Wunderlich *et al.* 2014; Gallego Romero *et al.* 2015; Ramaswamy *et al.* 2015) is warranted. iPSCs are an incredibly useful tool and provide a great deal of insight into questions of pluripotentiality and differentiation (Takahashi and Yamanaka 2006); however, some questions remain regarding whether epigenetic reprogramming during the iPSC protocol results in full pluripotency potential (Bilic and Belmonte 2012; Robinton and Daley 2012; Halevy and Urbach 2014). There are also known differences between iPSCs and ES cells in epigenetic landscape, transcribed genes, mutational load, and differentiation potential (Bilic and Belmonte 2012). For these reasons, we decided to take the more conservative path and use traditional ES cells to reduce potential confounding factors while testing our original hypotheses.

By characterizing the translational activity of the ancestral form of *NR2C1*, we were able to polarize the evolution of regulation of *Pepck* promoter activity. Specifically, amino acid substitutions occurring at different sites in humans and chimpanzees since their LCA increased transcriptional activities in both lineages. The substitutions that occurred along the chimpanzee lineage appear to have a greater effect on transcriptional activation than those that occurred along the

human lineage. In broad terms, this is an example of parallel evolution, as the phenotypic trajectory (increased transcriptional activation) was the same along both lineages. It is interesting that the MLE distribution for *NR2C1* suggested positive selection, as this implies that parallel evolution of gene expression could have been driven by Darwinian positive selection.

Acknowledgments

The work described here was supported by a National Science Foundation (NSF) doctoral dissertation research improvement grant (BSC-1455625), a Wenner-Gren Foundation dissertation fieldwork grant (8735), an NSF Integrative Graduate Education and Research traineeship grant (DGE-0801634), an NSF-Human Origins Moving in New Directions (HOMINID) grant (BCS-0827546), a grant from the James S. McDonnell Foundation (220020293), George Washington (GW) Office of the Vice President for Research and the GW School of Medicine for the GW Institute of Neuroscience Biomarker Core Facility, National Institutes of Health grant (DC-001534), Natural Sciences and Engineering Research Council of Canada (DG298394), Centre for Comparative Genomics and Evolutionary Bioinformatics (funded by the Tula Foundation), and Canadian Institutes of Health Research (CMF-108026).

Literature Cited

Alemseged, Z., F. Spoor, W. H. Kimbel, R. Bobe, D. Geraads *et al.*, 2006 A juvenile early hominin skeleton from Dikika, Ethiopia. Nature 443: 296–301.

Allman, J. M., N. A. Tetreault, A. Y. Hakeem, K. F. Manaye, K. Semendeferi *et al.*, 2010 The von Economo neurons in frontoinsular and anterior cingulate cortex in great apes and humans. Brain Struct. Funct. 214: 495–517.

Anderson, A. M., K. W. Carter, D. Anderson, and M. J. Wise, 2012 Coexpression of nuclear receptors and histone methylation modifying genes in the testis: implications for endocrine disruptor modes of action. PLoS One 7: e34158.

Anisimova, M., R. Nielsen, and Z. Yang, 2003 Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics 164: 1229–1236.

Baker, J. L., B. Wood, B. A. Karpinski, A.S. LaMantia, and T. M. Maynard, 2016 Testicular receptor 2, Nr2c1, is associated with stem cells in the developing olfactory epithelium and other cranial sensory and skeletal structures. Gene Expr. Patterns 20(1): 71–79.

Bao, L., H. Gu, K. A. Dunn, and J. P. Bielawski, 2007 Methods for selecting fixed-effect models for heterogeneous codon evolution, with comments on their application to gene and genome data. BMC Evol. Biol. 7: S5.

Benoit, G., A. Cooney, V. Giguere, H. Ingraham, M. Lazar *et al.*, 2006 International Union of Pharmacology. LXVI. Orphan nuclear receptors. Pharmacol. Rev. 58: 798–836.

Bielawski, J. P., and Z. Yang, 2004 A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J. Mol. Evol. 59: 121–132.

Bielawski, J. P., J. L. Baker, and J. Mingrone, 2016 Inference of episodic changes in natural selection acting on protein coding sequences via CODEML. Curr. Prot. Bioinf. (in press).

Bilic, J., and J. C. Izpisua Belmonte, 2012 Concise review: induced pluripotent stem cells *vs.* embryonic stem cells: Close enough or yet too far apart? Stem Cells 30: 33–41.

Bloch, N. I., J. M. Morrow, B. S. W. Chang, and T. D. Price, 2015 SWS2 visual pigment evolution as a test of historically contingent patterns of plumage color evolution in warblers. Evolution 69: 341–356.

Boddy, A. M., M. R. McGowen, C. C. Sherwood, L. I. Grossman, M. Goodman *et al.*, 2012 Comparative analysis of encephalization in mammals reveals relaxed constraints on anthropoid primate and cetacean brain scaling. J. Evol. Biol. 25: 981–994.

Boiani, M., 2002 Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes Dev. 16: 1209–1219.

Bourguet, W., P. Germain, and H. Gronemeyer, 2000 Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. Trends Pharmacol. Sci. 21(10): 381–388.

Braun, D. R., J. W. K. Harris, N. E. Levin, J. T. McCoy, A. I. R. Herries *et al.*, 2010 Early hominin diet included diverse terrestrial and aquatic animals 1.95 Ma in East Turkana, Kenya. Proc. Natl. Acad. Sci. USA 107: 10002–10007.

Brayer, K. J., V. J. Lynch, and G. P. Wagner, 2011 Evolution of a derived protein-protein interaction between HoxA11 and Foxo1a in mammals caused by changes in intramolecular regulation. Proc. Natl. Acad. Sci. USA 108: E414–E420.

Bridgham, J. T., E. A. Ortlund, and J. W. Thornton, 2009 An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. Nature 461: 515–519.

Chambers, I., and A. Smith, 2004 Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene 23: 7150–7160.

Chang, B. S. W., K. Jonsson, M. A. Kazmi, M. J. Donoghue, and T. P. Sakmar, 2002 Recreating a functional ancestral archosaur visual pigment. Mol. Biol. Evol. 19: 1483–1489.

Chang, C., and J. Kokontis, 1988 Identification of a new member of the steroid receptor super-family by cloning and sequence analysis. Biochem. Biophys. Res. Commun. 155: 971–977.

Chen, C., J. C. Opazo, O. Erez, M. Uddin, J. Santolaya-Forgas *et al.*, 2008 The human progesterone receptor shows evidence of adaptive evolution associated with its ability to act as a transcription factor. Mol. Phylogenet. Evol. 47: 637–649.

Eick, G. N., and J. W. Thornton, 2011 Evolution of steroid receptors from an estrogen-sensitive ancestral receptor. Mol. Cell. Endocrinol. 334: 31–38.

Enmark, E., and J. A. Gustafsson, 1996 Orphan nuclear receptors: the first eight years. Mol. Endocrinol. 10: 1293–1307.

Fletcher, W., and Z. Yang, 2010 The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. Mol. Biol. Evol. 27: 2257–2267.

Forsberg, R., and F. B. Christiansen, 2003 A codon-based model of host-specific selection in parasites, with an application to the influenza A virus. Mol. Biol. Evol. 20: 1252–1259.

Gallego Romero, I., B. J. Pavlovic, I. Hernando-Herraez, X. Zhou, M. C. Ward *et al.*, 2015 A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics. eLife 4: e07103.

Gaucher, E. A., J. M. Thomson, M. F. Burgan, and S. A. Benner, 2003 Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 425: 285–288.

Gaucher, E. A., S. Govindarajan, and O. K. Ganesh, 2008 Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature 451: 704–707.

Germain, P., B. Staels, C. Dacquet, M. Spedding, and V. Laudet, 2006 Overview of nomenclature of nuclear receptors. Pharmacol. Rev. 58: 685–704.

Goldman, N., and Z. Yang, 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11: 725–736.

Gruppuso, P. A., T. C. Bienieki, and R. A. Faris, 1999 The relationship between differentiation and proliferation in late gestation fetal rat hepatocytes. Pediatr. Res. 46: 14–19.

Halevy, T., and A. Urbach, 2014 Comparing ESC and iPSC-based models for human genetic disorders. J. Clin. Med. 3(4): 1146–1162.

Harms, M. J., and J. W. Thornton, 2010 Analyzing protein structure and function using ancestral gene reconstruction. Curr. Opin. Struct. Biol. 20: 360–366.

Harms, M. J., and J. W. Thornton, 2013 Evolutionary biochemistry: revealing the historical and physical causes of protein properties. Nat. Rev. Genet. 14: 559–571.

Hay, D. C., L. Sutherland, J. Clark, and T. Burdon, 2004 Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. Stem Cells 22: 225–235.

Hu, Y.-C., C.-R. Shyr, W. Che, X.-M. Mu, E. Kim *et al.*, 2002 Suppression of estrogen receptor-mediated transcription and cell growth by interaction with TR2 orphan receptor. J. Biol. Chem. 277: 33571–33579.

Jungers, W. L., W. E. H. Harcourt-Smith, R. E. Wunderlich, M. W. Tocheri, S. G. Larson *et al.*, 2009a The foot of Homo floresiensis. Nature 459: 81–84.

Jungers, W. L., S. G. Larson, W. Harcourt-Smith, M. J. Morwood, T. Sutikna *et al.*, 2009b Descriptions of the lower limb skeleton of Homo floresiensis. J. Hum. Evol. 57: 538–554.

Katoh, K., 2002 MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059–3066.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle *et al.*, 2002 The Human Genome Browser at UCSC. Genome Res. 12: 996–1006.

Ketterson, E. D., J. W. Atwell, and J. W. McGlothlin, 2009 Phenotypic integration and independence: hormones, performance, and response to environmental change. Integr. Comp. Biol. 49: 365–379.

Kohn, J. A., K. Deshpande, and E. A. Ortlund, 2012 Deciphering modern glucocorticoid cross-pharmacology using ancestral corticosteroid receptors. J. Biol. Chem. 287: 16267–16275.

Kosakovsky Pond, S. L., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost, 2006 GARD: a genetic algorithm for recombination detection. Bioinformatics 22: 3096–3098.

Krasowski, M. D., K. Yasuda, L. R. Hagey, and E. G. Schuetz, 2005 Evolutionary selection across the nuclear hormone receptor superfamily with a focus on the NR1I subfamily (vitamin D, pregnane X, and constitutive androstane receptors). Nucl. Recept. 3: 2.

Kratzer, J. T., M. A. Lanaspa, M. N. Murphy, C. Cicerchi, C. L. Graves *et al.*, 2014 Evolutionary history and metabolic insights of ancient mammalian uricases. Proc. Natl. Acad. Sci. USA 111: 3763–3768.

Laudet, V., 1997 Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. J. Mol. Endocrinol. 19: 207–226.

Lee, H. J., and C. Chang, 1995 Identification of human TR2 orphan receptor response element in the transcriptional initiation site of the simian virus 40 major late promoter. J. Biol. Chem. 270: 5434–5440.

Lee, Y.-F., H.-J. Lee, and C. Chang, 2002 Recent advances in the TR2 and TR4 orphan receptors of the nuclear receptor superfamily. J. Steroid Biochem. Mol. Biol. 81: 291–308.

Liu, Y., X. S. Liu, L. Wei, R. B. Altman, and S. Batzoglou, 2004 Eukaryotic regulatory element conservation analysis and identification using comparative genomics. Genome Res. 14: 451–458.

Lucas, P. C., R. M. O'Brien, J. A. Mitchell, C. M. Davis, E. Imai *et al.*, 1991 A retinoic acid response element is part of a pleiotropic domain in the phosphoenolpyruvate carboxykinase gene. Proc. Natl. Acad. Sci. USA 88: 2184–2188.

Mandal, S., A. G. Lindgren, A. S. Srivastava, A. T. Clark, and U. Banerjee, 2011 Mitochondrial function controls proliferation and early differentiation potential of embryonic stem cells. Stem Cells 29: 486–495.

Marchetto, M. C., I. Narvaiza, A. M. Denli, C. Benner, T. A. Lazzarini *et al.*, 2013 Differential L1 regulation in pluripotent stem cells of humans and apes. Nature 503: 525–529.

Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire *et al.*, 2011 CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39: D225–D229.

Maynard, T., D. Gopalakrishna, D. Meechan, E. Paronett, J. Newbern *et al.*, 2013 22q11 Gene dosage establishes an adaptive range for sonic hedgehog and retinoic acid signaling during early development. Hum. Mol. Genet. 22(2): 300–312.

Mcbrearty, S., and A. S. Brooks, 2000 The revolution that wasn't: a new interpretation of the origin of modern human behavior. J. Hum. Evol. 39: 453–563.

Mitsui, K., Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami *et al.*, 2003 The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113: 631–642.

Muse, S. V., and B. S. Gaut, 1994 A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11: 715–724.

Nichols, J., B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius *et al.*, 1998 Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95: 379–391.

Niwa, H., J. Miyazaki, and A. G. Smith, 2000 Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. 24: 372–376.

Ohi, Y., H. Qin, C. Hong, L. Blouin, J. M. Polo *et al.*, 2011 Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. Nat. Cell Biol. 13(5): 541–549.

Pikarsky, E., H. Sharir, E. Ben-Shushan, and Y. Bergman, 1994 Retinoic acid represses Oct-3/4 gene expression through several retinoic acid-responsive elements located in the promoter-enhancer region. Mol. Cell. Biol. 14: 1026–1038.

Ramaswamy, K., W. Y. Yik, X. M. Wang, E. N. Oliphant, W. Lu *et al.*, 2015 Derivation of induced pluripotent stem cells from orangutan skin fibroblasts. BMC Res. Notes 16(8): 577.

Robinson-Rechavi, M., A.-S. Carpentier, M. Duffraisse, and V. Laudet, 2001 How many nuclear hormone receptors are there in the human genome? Trends Genet. 17: 554–556.

Robinton, D. A., and G. Q. Daley, 2012 The promise of induced pluripotent stem cells in research and therapy. Nature 481: 295–305.

Rubinstein, N. D., A. Doron-Faigenboim, I. Mayrose, and T. Pupko, 2011 Evolutionary models accounting for layers of selection in protein-coding genes and their impact on the inference of positive selection. Mol. Biol. Evol. 28(12): 3297–3308.

Roesler, W. J., G. R. Vandenbark, and R. W. Hanson, 1989 Identification of multiple protein binding domains in the promoter-regulatory region of the phosphoenolpyruvate carboxykinase (GTP) gene. J. Biol. Chem. 264(16): 9657–9664.

Schneider, A., A. Souvorov, N. Sabath, G. Landan, G. H. Gonnet *et al.*, 2009 Estimates of positive Darwinian selection are inflated by errors in sequencing, annotation, and alignment. Genome Biol. Evol. 1: 114–118.

Schott, R. K., S. P. Refvik, F. E. Hauser, H. López-Fernández, and B. S. W. Chang, 2014 Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. Mol. Biol. Evol. 31: 1149–1165.

Self, S. G., and K. Y. Liang, 1987 Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Amer. Stat. Assoc. 82(398): 605–610.

Sherwood, C. C., and T. Duka, 2012 Now that we've got the map, where are we going moving from gene candidate lists to function in studies of brain evolution. Brain Behav. Evol. 80: 167–169.

Shyr, C.-R., H.-Y. Kang, M.-Y. Tsai, N.-C. Liu, P.-Y. Ku *et al.*, 2009 Roles of testicular orphan nuclear receptors 2 and 4 in early embryonic development and embryonic stem cells. Endocrinology 150: 2454–2462.

Stamatakis, A., 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

Storey, J. D., 2002 A direct approach to false discovery rates. J. R. Stat. Soc. B 64: 479–498.

Takahashi, K., and S. Yamanaka, 2006 Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126: 663–676.

Thornton, J. W., 2001 Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc. Natl. Acad. Sci. USA 98: 5671–5676.

Thornton, J. W., E. Need, and D. Crews, 2003 Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 301: 1714–1717.

Tryon, C. A., N. T. Roach, and M. A. V. Logan, 2008 The Middle Stone Age of the northern Kenyan Rift: age and context of new archaeological sites from the Kapedo Tuffs. J. Hum. Evol. 55: 652–664.

Ugalde, J. A., B. S. W. Chang, and M. V. Matz, 2004 Evolution of coral pigments recreated. Science 305: 1433.

Ward, C. V., W. H. Kimbel, and D. C. Johanson, 2011 Complete fourth metatarsal and arches in the foot of Australopithecus afarensis. Science 331: 750–753.

Wasserman, W. W., M. Palumbo, W. Thompson, J. W. Fickett, and C. E. Lawrence, 2000 Human-mouse genome comparisons to locate regulatory sites. Nat. Genet. 26: 225–228.

Weadick, C. J., and B. S. W. Chang, 2012 An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. Mol. Biol. Evol. 29: 1297–1300.

Whelan, S., and N. Goldman, 2001 A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18: 691–699.

Williamson, S. H., M. J. Hubisz, A. G. Clark, B. A. Payseur, C. D. Bustamante *et al.*, 2007 Localizing recent adaptive evolution in the human genome. PLoS Genet. 3: e90.

Wood, B., 1996 Hominid palaeobiology: Have studies of comparative development come of age? Am. J. Phys. Anthropol. 99: 9–15.

Wunderlich, S., M. Kircher, B. Vieth, A. Haase, S. Merkert *et al.*, 2014 Primate iPS cells as tools for evolutionary analyses. Stem Cell Res. (Amst.) 12: 622–629.

Yang, Z., 1994 Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39: 105–111.

Yang, Z., 2007 PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24: 1586–1591.

Yang, Z., and R. Nielsen, 2002 Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. 19: 908–917.

Yang, Z., and W. J. Swanson, 2002 Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. Mol. Biol. Evol. 19: 49–57.

Zaehres, H., M. W. Lensch, L. Daheron, S. A. Stewart, J. Itskovitz-Eldor *et al.*, 2005 High-efficiency RNA interference in human embryonic stem cells. Stem Cells 23: 299–305.

Zhang, J., R. Nielsen, and Z. Yang, 2005 Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22: 2472–2479.

Zimmer, D. B., and M. A. Magnuson, 1990 Immunohistochemical localization of phosphoenolpyruvate carboxykinase in adult and developing mouse tissues. J. Histochem. Cytochem. 38: 171–178.

Communicating editor: J. M. Akey

GENETICS

Supporting Information
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.183889/-/DC1

Functional Divergence of the Nuclear Receptor *NR2C1* as a Modulator of Pluripotentiality During Hominid Evolution

Jennifer L. Baker, Katherine A. Dunn, Joseph Mingrone, Bernard A. Wood, Beverly A. Karpinski, Chet C. Sherwood, Derek E. Wildman, Thomas M. Maynard, and Joseph P. Bielawski

DOI: 10.1534/genetics.115.183889

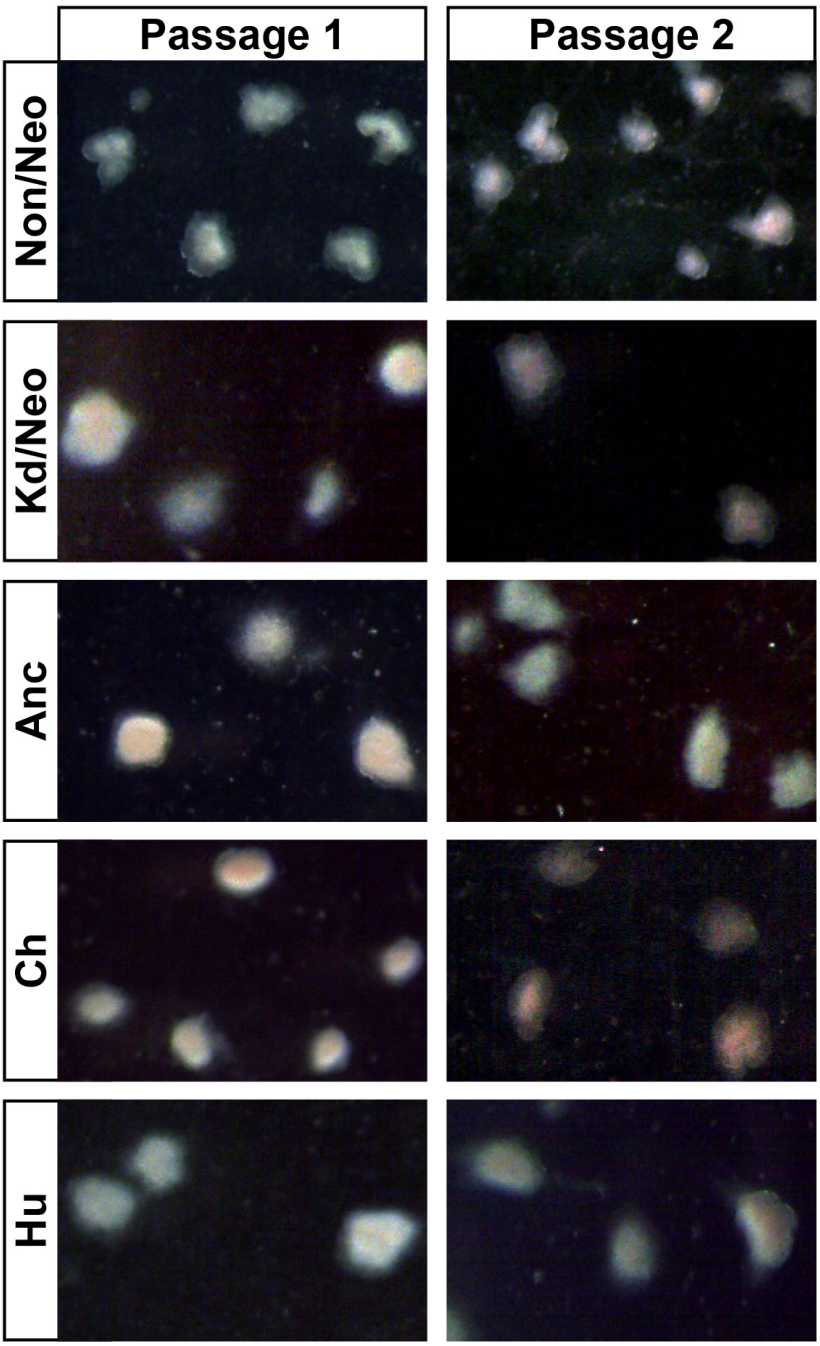

Figure S1. Brightfield images of two passages for the five sets of undifferentiated ES colonies, demonstrating they are rounded colonies. Key: Hu (human), Ch (chimpanzee), Anc (ancestral), Kd (knockdown), Non (nonsense)

Table S1. Summary of the 146 NR sequences initially downloaded from the UCSC Genome Browser from which 48 sequences were selected for an evolutionary survey in Phase 2 of this study.

Gene Name	NM #	Variant IDs	Notes:
AR	NM_000044	AR-i1	1
AR	NM_001011645	AR-i2	
ESR1	NM_000125	ESR1	1
ESR2	NM_001437	ESR2-i1	1
ESR2	NM_001040275	ESR2-i2	
ESR2	NM_001214902	ESR2-i3	
ESR2	NM_001271876	ESR2-i5	
ESR2	NM_001214903	ESR2	2
ESR2	NM_001271877	ESR2-i6	
ESRRA	NM_001282450	ESRRA-i1	1, 3
ESRRA	NM_001282451	ESRRA-i2	
ESRRB	NM_004452	ESRRB	1, 3, 4
ESRRG	NM_001243518	ESRRG-i6	
ESRRG	NM_001438	ESRRG-i1	1
ESRRG	NM_001134285	ESRRG-i2	
ESRRG	NM_001243507	ESRRG-i5	
ESRRG	NM_001243506	ESRRG-i4	
ESRRG	NM_001243505	ESRRG-i3	
HNF4A	NM_000457	HNF4A-α2	1
HNF4A	NM_001258355	HNF4A-α4	
HNF4A	NM_178849	HNF4A-α1	
HNF4A	NM_175914	HNF4A-α8	
HNF4A	NM_001287183	HNF4A-α11	
HNF4A	NM_001030003	HNF4A-α7	
HNF4A	NM_001287182	HNF4A-α10	
HNF4A	NM_178850	HNF4A-α3	
HNF4A	NM_001030004	HNF4A-α9	
HNF4A	NM_001287184	HNF4A-α12	
HNF4G	NM_004133	HNF4G	1
NR0B1	NM_000475	NR0B1	1
NR0B2	NM_021969	NR0B2	1
NR1D1	NM_021724	NR1D1	1, 3
NR1D2	NM_005126	NR1D2-i1	1
NR1D2	NM_001145425	NR1D2-i2	
NR1H2	NM_007121	NR1H2	1
NR1H2	NM_001256647	NR1H2	
NR1H3	NM_001251934	NR1H3-i4	
NR1H3	NM_005693	NR1H3-i1	1
NR1H3	NM_001130102	NR1H3-i3	
NR1H3	NM_001130101	NR1H3-i2	
NR1H4	NM_001206993	NR1H4-i3	1

NR1H4	NM_001206992	NR1H4-i4	
NR1H4	NM_001206977	NR1H4-i1	
NR1H4	NM_005123	NR1H4-i2	
NR1H4	NM_001206978	NR1H4-i5	
NR1I2	NM_022002	NR1I2-i2	1
NR1I2	NM_003889	NR1I2-i1	
NR1I2	NM_033013	NR1I2-i3	
NR1I3	NM_001077482	NR1I3-i1	
NR1I3	NM_001077480	NR1I3-i2	
NR1I3	NM_005122	NR1I3-i3	1
NR1I3	NM_001077469	NR1I3-i6	
NR1I3	NM_001077478	NR1I3-i7	
NR1I3	NM_001077472	NR1I3-i9	
NR1I3	NM_001077479	NR1I3-i10	
NR1I3	NM_001077473	NR1I3-i12	
NR1I3	NM_001077481	NR1I3-i4	
NR1I3	NM_001077476	NR1I3-i13	
NR1I3	NM_001077471	NR1I3-i5	
NR1I3	NM_001077477	NR1I3-i14	
NR1I3	NM_001077474	NR1I3-i8	
NR1I3	NM_001077470	NR1I3-i11	
NR1I3	NM_001077475	NR1I3-i15	
NR2C1	NM_003297	NR2C1-ia	1, 3, 4
NR2C1	NM_001127362	NR2C1-ic	
NR2C1	NM_001032287	NR2C1-ib	
NR2C2	NM_003298	NR2C2	1
NR2E1	NM_001286102	NR2E1-ia	
NR2E1	NM_003269	NR2E1-ib	1
NR2E3	NM_014249	NR2E3-ib	1, 3
NR2E3	NM_016346	NR2E3-ia	
NR2F1	NM_005654	NR2F1	1
NR2F2	NM_021005	NR2F2-ia	1
NR2F2	NM_001145155	NR2F2-ib	
NR2F2	NM_001145156	NR2F2-ic	
NR2F6	NM_005234	NR2F6	1
NR3C1	NM_001024094	NR3C1-g	1
NR3C1	NM_000176	NR3C1-α	
NR3C1	NM_001204258	NR3C1-αB	
NR3C1	NM_001020825	NR3C1-b	
NR3C1	NM_001204259	NR3C1-αC1	
NR3C1	NM_001204260	NR3C1-αC2	
NR3C1	NM_001204261	NR3C1-αC3	
NR3C1	NM_001204262	NR3C1-αD1	
NR3C1	NM_001204263	NR3C1-αD2	
NR3C1	NM_001204264	NR3C1-αD3	
NR3C1	NM_001204265	NR3C1 GR-P	

NR3C2	NM_000901	NR3C2-i1	1
NR3C2	NM_001166104	NR3C2-i2	
NR4A1	NM_001202233	NR4A1-i2	1
NR4A1	NM_002135	NR4A1-i1	
NR4A2	NM_006186	NR4A2	1
NR4A3	NM_173200	NR4A3-ib	1
NR4A3	NM_006981	NR4A3-ia	
NR4A3	NM_173199	NR4A3-ic	
NR5A1	NM_004959	NR5A1	1
NR5A2	NM_205860	NR5A2-i1	1
NR5A2	NM_003822	NR5A2-i2	
NR5A2	NM_001276464	NR5A2-i3	
NR6A1	NM_033334	NR6A1-i1	1
NR6A1	NM_001278546	NR6A1-i4	
NR6A1	NM_001489	NR6A1-i2	
PGR	NM_000926	PGR-iB	
PGR	NM_001202474	PGR-iA	
PGR	NM_001271161	PGR-iC	
PGR	NM_001271162	PGR-iD	1, 3, 4
PPARA	NM_001001928	PPARA	1
PPARD	NM_001171818	PPARD-i1	1
PPARD	NM_001171819	PPARD-i3	
PPARD	NM_177435	PPARD-i2	
PPARD	NM_001171820	PPARD-i4	
PPARG	NM_015869	PPARG-i2	1, 3
PPARG	NM_138711	PPARG-i1	
RARA	NM_000964	RARA-i1	1
RARA	NM_001145301	RARA-i1	
RARA	NM_001024809	RARA-i2	
RARA	NM_001145302	RARA-i4	
RARB	NM_001290216	RARB-i3	1
RARB	NM_000965	RARB-i1	
RARB	NM_001290300	RARB-i6	
RARB	NM_001290277	RARB-i5	
RARB	NM_001290266	RARB-i4	
RARB	NM_001290217	RARB-i2	
RARG	NM_000966	RARG-i1	1
RARG	NM_001042728	RARG-i2	
RARG	NM_001243732	RARG-i3	
RARG	NM_001243730	RARG-i4	
RARG	NM_001243731	RARG-i5	
RORA	NM_134260	RORA-ib	
RORA	NM_002943	RORA-ic	
RORA	NM_134261	RORA-ia	1, 3, 4
RORA	NM_134262	RORA-id	
RORB	NM_006914	RORB	1

RORC	NM_005060	RORC-ia	1
RORC	NM_001001523	RORC-ib	
RXRA	NM_002957	RXRA	1
RXRB	NM_001270401	RXRB-t1	1
RXRB	NM_021976	RXRB-t2	
RXRG	NM_006917	RXRG-ia	1
RXRG	NM_001256570	RXRG-ic	
THRA	NM_001190919	THRA-i2	1
THRA	NM_001190918	THRA-i3	
THRA	NM_199334	THRA-i1	
THRB	NM_000461	THRB	1
VDR	NM_001017536	VDR-iVDRB1	
VDR	NM_000376	VDR	1

Notes:

1: Indicates each of the 48 genes that was selected for evolutionary modeling and analysis in this study

2: Indicates that the RefSeq was permanently suppressed after the data we downloaded because it represents a partial transcript sequence. The complete transcript was a nonsense-mediated mRNA decay (NMD) candidate.

3: Indicates each of nine initial candidate genes identified in Phase 2 of in this study

4: Indicates each of the final four candidates selected at the end of Phase 2 as warranting further investigation

File S1: Assessment of among-domain evolutionary heterogeneity within NRs via fixed-effect codon models

NRs are routinely divided into five structural domains. We used the Conserved Domain Database (Marchler-Bauer et al., 2011) to map each aligned site to four of the five major domains: (*i*) the NTD, (*ii*) the DBD, (*iii*) the flexible hinge domain, and (*iv*) the LBD. The fifth domain (CTD) was excluded from this analysis, as it was not present in every NR. We modeled heterogeneous evolution among domains by using codon models, where independent evolutionary categories were *a priori* specified within a codon model for each of the four domains given above (Yang and Swanson, 2002; Bao et al., 2007). In this way the evolution of each site was modeled as evolving according to the structural domain to which it had been mapped (*i.e.*, it was treated as a fixed effect [FE] in the codon model). The following four evolutionary processes were permitted to differ among the structural partitions: selection intensity (ω), overall rate of evolution via a branch-length scale parameter (c), transition-transversion ratio (κ), and equilibrium codon frequencies (π_j). This permits 16 different combinations of among-domain heterogeneity and homogeneity, with the simplest model (FE16) specific complete homogeneity and the most complex model (FE1) specifying complete heterogeneity. The full set of 16 models are:

Model ID:	Among domain heterogeneity:
FE1	κ, ω, c, π
FE2	κ, ω, c
FE3	κ, ω, π
FE4	κ, ω
FE5	κ, c, π
FE6	κ, c
FE7	κ, π
FE8	κ
FE9	ω, c, π
FE10	ω, c
FE11	ω, π
FE12	ω
FE13	c, π
FE14	c
FE15	π
FE16	[none]

The best-fit model for each gene was selected according to a forward addition procedure based on likelihood ratio tests (Bao et al., 2007). Starting with the most complex model (FE1) parameters are sequentially added to the model that provide a significant improvement to the fit of the model. Because the likelihood ratio does not follow the expected χ^2 distribution (Aagaard and Phillips, 2005; Bao et al., 2007), significance was determined by using likelihood ratio tests with an adjusted cut-off (the threshold for significance was 0.0001 rather than 0.05), as recommend by Bao et al. (2007). We do not use AIC or AICc, as the potential for different degrees of non-independence among partitions within the FE models means that AIC is not likely to be a good approximation to the Kullback-Leibler divergence. Fitting the model requires that each site within an alignment is unambiguously assigned to a structural category within the codon model; these site assignments are included in the files deposited in the DRYAD data repository *(*doi:10.5061/dryad.bg3g3). Likelihood calculations and parameter estimates were obtained by using a modified version of the codeml program from of the PAML package called codeml_FE (available from the authors upon request).

The best-fit model for each gene is given in **Table S1.1** below. Based on these results, the most common source of among-domain variation was ω (36 genes in total), which was expected based on Figure 4 in the main text. The next most common factor was the branch length scale parameter (25 genes). Codon frequencies were a minor factor (13 genes), and exhibited only a small effect size. No genes had models that included variable κ parameters.

References

Aagaard J.E., Phillips P., (2005). Accuracy and power of the likelihood ratio test for comparing evolutionary rates among genes. *J. Mol. Evol*. 60:426-33.

Bao, L., Gu, H., Dunn, K.A., Bielawski, J.P., (2007). Methods for selecting fixed-effect models for heterogeneous codon evolution, with comments on their application to gene and genome data. *BMC Evol. Biol*. 7:Suppl. 1, S5.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res*. 39:D225–9.

Yang, Z., Swanson, W.J., (2002). Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol*. 19:49-57.

Table S1.1 Best-fit fixed-effect (FE) codon model for heterogeneous evolution among the structural domains of NRs

Gene name	Model	Gene name	Model
AR	FE10	*NR3C2*	FE9
ESR1	FE9	*NR4A1*	FE12
ESR2	FE9	*NR4A2*	FE15
ESRRA	FE12	*NR4A3*	FE9
ESRRB[4]	FE15	*NR5A1*	FE12
ESRRG	FE14	*NR5A2*	FE10
HNF4A	FE16	*NR6A1*	FE12
HNF4G	FE12	***PGR***[3]	FE9
NR1D1	FE9	*PPARA*	FE9
NR1D2	FE10	*PPARD*	FE12
NR1H2	FE10	***PPARG***	FE9
NR1H3	FE9	*RARA*	FE16
NR1H4	FE12	*RARB*	FE14
NR1I2	FE10	***RARG***	FE12
NR1I3	FE16	***RORA***[2]	FE9
NR2C1[1]	FE12	*RORB*	FE14
NR2C2	FE12	*RORC*	FE10
NR2E1	FE13	*RXRA*	FE12
NR2E3	FE12	*RXRB*	FE11
NR2F1	FE10	*RXRG*	FE12
NR2F2	FE13	*THRA*	FE10
NR2F6	FE12	*THRB*	FE10
NR3C1	FE12	*VDR*	FE10

Note: The complexity of each model (FE1-FE16) is given above. The nine initial candidate genes identified in Phase 2 of this study are shown in bold. The four final candidates selected at the end of Phase 2 as warranting further investigation are numbered.

File S2: Overview of codon models and likelihood ratio tests for episodic and selective shifts.

1. The basic model. Codon models provide a powerful means of detecting changes in the intensity of natural selection according to the relative rate of non-synonymous and synonymous substitutions (Bielawski and Yang, 2005). The general framework is the Markov model, and such models were independently developed for codon evolution by Goldman and Yang (1994) and Muse and Gaut (1994). A Markovian process is a simple stochastic process used to describe changes between states, and codon models use it to describe the process of substitution between the sense codons of a protein-coding sequence. The stop codons are excluded from the process because they are not expected to occur at any site other than the last codon of a sequence that encodes a functional protein.

The models used in this study are derived from the basic model of Goldman and Yang (1994), which is very similar to the model developed by Muse and Gaut (1994). The basic form of this model is specified by the instantaneous rate matrix $Q = \{q_{ij}\}$, which gives the rate of substitution rate from any initial codon i to another codon j.

$$q_{ij} = \begin{cases} 0, & \text{if } i \text{ and } j \text{ differ at two or 3 codon positions,} \\ \mu\pi_j, & \text{if } i \text{ and } j \text{ differ by a synonymous transition,} \\ \mu\kappa\pi_j, & \text{if } i \text{ and } j \text{ differ by a synonymous transversion,} \\ \mu\omega\pi_j, & \text{if } i \text{ and } j \text{ differ by a nonsynonymous transition,} \\ \mu\kappa\omega\pi_j, & \text{if } i \text{ and } j \text{ differ by a nonsynonymous transversion.} \end{cases}$$

Because genes often have unequal codon usage, the rate of change between codons i and j (parameter μ) is multiplied by the parameter π_j, which is the equilibrium frequency of the j^{th} codon. The process of evolution at the DNA level can be heterogeneous. This model accommodates DNA-level processes by multiplying the rate by the parameter κ (the transition/transversion rate ratio) when the change between codons involves a nucleotide transition. The κ parameter corrects for the widely observed substitution rate difference between transitions and transversions. Since our primary interest is to use the d_N/d_S ratio to quantify the intensity of selection pressure at the level of the protein (Bielawski and Yang 2005), this rate ratio is inlcuded in the model as $\omega = d_N/d_S$. Specifically, when the change between codons is nonsynonymous, the rate is multiplied by ω to adjust it for the effect of

selection acting on the protein-product of the gene. Purifying selection lowers the relative rate of nonsynonymous substitution, whereas positive diversifying selection is expected to raise it.

The diagonal elements of the matrix Q are defined by the requirement that the rows of the matrix sum to zero. The elements of the matrix are used to obtain $p_{ij}(t)$, which is the probability that codon i is substituted with codon j over time t. The values of parameters κ, ω and t are estimated from the data "in hand" via maximum likelihood. A review of maximum likelihood estimation for codon models is provided in Bielawski and Yang (2005). Because separate estimation of μ and t is not possible, the rate μ is fixed such that the expected number of substitutions per codon will be equal to 1. This scaling means that the branch lengths of a phylogeny (t_i) are measured in terms of the mean number of substitutions per codon.

2. Models for variable selection pressure among sites. There are a wide variety of models that permit the parameter ω to vary among sites. When there is insufficient knowledge to partition codon sites into different classes of selection pressure (having different ω parameters), a statistical distribution is used to model ω variation among sites. In this study we employ two such codon models (referred to as M2a and M3). We illustrate the approach with the "discrete model" (M3) of Yang et al. 2000. M3 divides codon sites into K discrete classes (i = 0, 1, 2, … , K-1), with the $\omega = d_N/d_S$ ratios, and their proportion of sites within a gene, given as follows:

$$\omega_0, \omega_1 \ldots \omega_{K-1}$$
$$p_0, p_1 \ldots p_{K-1}$$

These are parameters of the ω distribution for M3. In this study we only used models with K=2 and K=3 for the ω distribution. Because we do not know which of the K classes a given site, h, will belong, the likelihood of the data at a site (x_h) is summed over the different categories for ω as follows:

$$f(x_h) = \sum_{i=0}^{K-1} p_i \, f(x_h|\omega_i)$$

The conditional probability that a site evolved according to the i^{th} class within the distribution is given above by $f(x_h|\omega_i)$. The unconditional probability of the data at a site, $f(x_h)$, can be thought of as an average over the site classes of the ω

distribution, and is computed as a weighted sum of the conditional probabilities. Assuming that the substitution process at one codon site is independent of the process at the other sites, the log of these site likelihoods can be summed over the *n* sites of the dataset.

$$\ell = \sum_{h=1}^{n} log\{f(x_h)\}$$

Maximum likelihood estimation of the parameters of the ω distribution can now be estimated by finding the values that maximize the log-likelihood of a multi-sequence alignment. The likelihood calculations for a multi-sequence alignment are carried out on a phylogeny. For an explanation of the likelihood, readers are referred to Bielawski and Yang (2005).

The ω parameters of model M3 are unconstrained, and the p_i parameters are constrained only so that their sum is equal to 1. Codon model M2a (Nielsen and Yang 1998) is obtained by setting K=3 and putting constraints on the ω parameters. Specifically one category is constrained to have purifying selection (ω_0<1), another is forced to have neutral evolution (ω_1=1), and the last is constrained to have positive selection (ω_2>1). Figure S2.1 below provides a graphical representation of codon models M2a and M3. Both models are useful for investigating among site variation in selection pressure, but in this study they serve as null models in LRTs for episodic and long-terms shifts in selection pressure (see section 4 below).

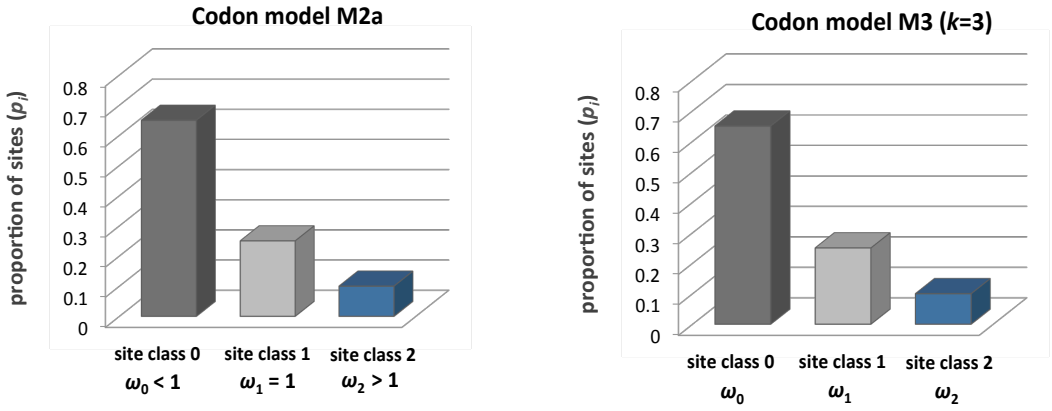

Figure S2.1. A graphical representation of the parameters of the ω distribution for codon models M2a and M3. Note that these models differ according the restrictions placed on the values that the ω parameters can take.

3. Models for variable selection pressure over time and among sites. The branch-site models (Model A and Model B) and clade-site models (Model C and Model D) are used to infer when the intensity of selection changes over evolutionary history. Because some sites will be essential to the core function of the encoded protein, selection intensity is expected to remain constant at those sites and vary over the tree at the other sites (Figures 2.2 and 2.3). Models A and B (Yang and Nielsen 2002; Zhang et al. 2005) are used to permit some sites to experience episodic (short-term) changes in selection intensity, and Models C and D (Bielawski and Yang 2004) are used to permit some sites to experience a long-term shift in the intensity of selection. All of these models require the user to specify *a priori* which branches have experienced a change in selection pressure; consequently, among-branch ω variation is a fixed-effect within these models. In contrast, they model among-site variation in selection pressure as a random effect (as do the sites-models described above) by using a statistical distribution.

Branch-site models: Branch-site codon models, like the sites-models, use a statistical distribution with several discrete site-classes (Figure 2.2A,B) for ω. Those site classes represent two broad evolutionary regimes: constant and episodic selection (Figure 2.2A). The constant selection regime is comprised of two site classes, where the same ω parameter is applied to all the branches of the phylogeny within a site class. The intensity of selection is permitted to differ among sites by using a different ω for each site class (ω_0 and ω_1). Within the episodic selection regime, sites will experience unique selection intensity along a particular branch, or branches, of a tree. Such branches are referred to as "foreground" (FG) branches, and the model employs an independent parameter for the selection intensity unique to the FG branch (ω_{FG}). The fraction of sites within a gene having unique selection along the FG branch is modeled by parameter p_{FG} (in Figure 2.2, $p_{FG} = p_{2a} + p_{2b}$). Note that selection intensity for all other branches is modeled using either ω_0 or ω_1; thus, those two parameters are shared among the constant and episodic regimes.

Branch-site Models A and B have the same basic structure (Figure S2.2A, B), differing only in the restrictions placed on the values of their ω parameters (Figure S2.2C). Model A is obtained by restricting the site-classes of the constant regime such that one of them must have purifying selection ($\omega_0 < 1$) and the other must have neutral evolution ($\omega_1 = 1$). For the episodic regime of Model A, one site-class permits episodic positive selection ($\omega_{FG}>1$) within a background of purifying selection ($\omega_0 < 1$) and the other permits episodic positive selection ($\omega_{FG}>1$) within a background of selectively neutral evolution ($\omega_1 = 1$). Model B places no restriction

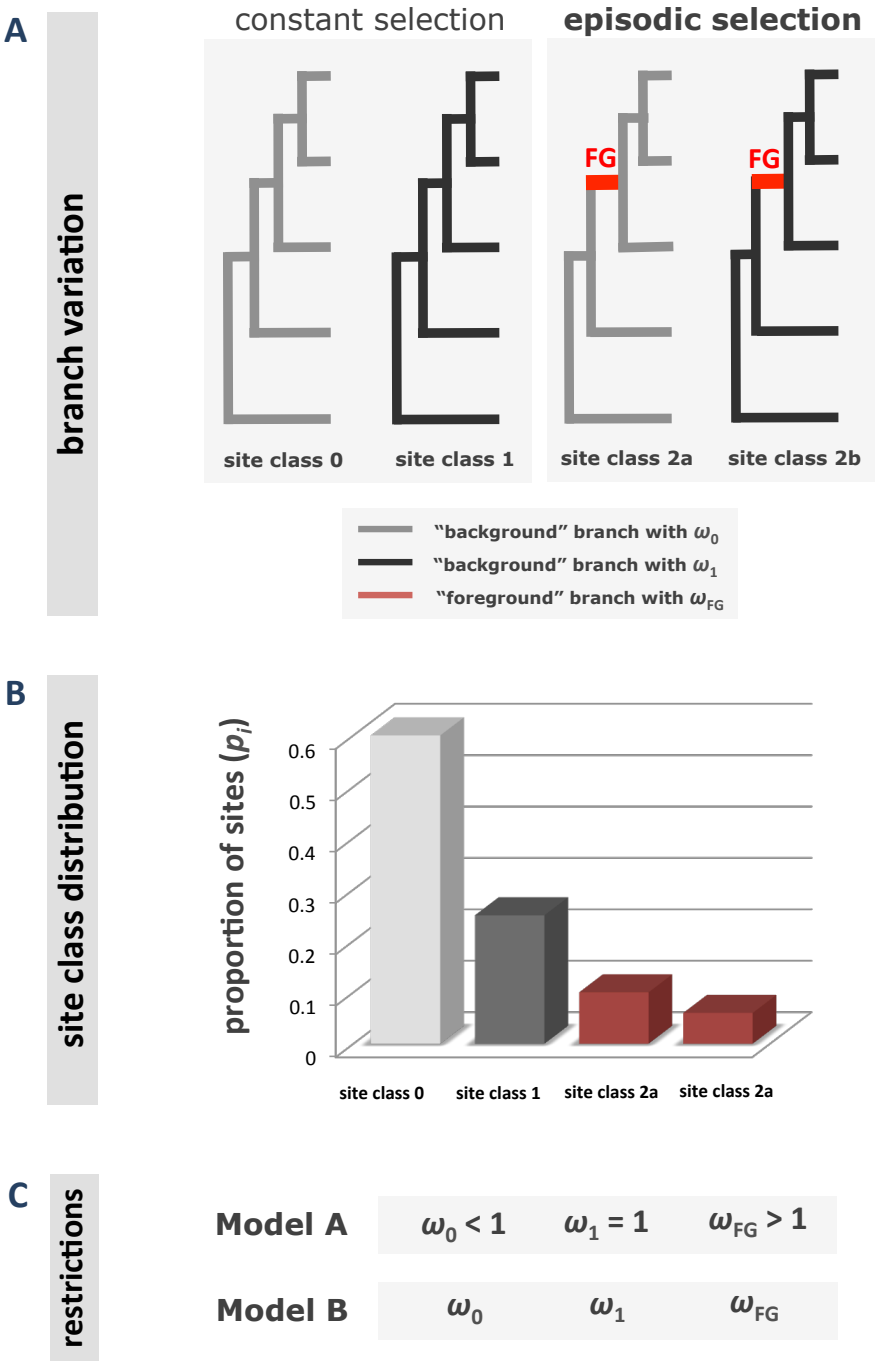

Figure S2.2. Graphical representation of how the ω parameters of branch-site codon models differ among branches of a phylogenetic tree and among site classes. (**A**) Two site classes (0 and 1) specify constant selection pressure over the tree, and two classes (2a and 2b) specify an episodic change in selection pressure along the foreground (FG) branch (**B**) The unconstrained discrete distribution for the four site classes of the codon model. (**C**) Models A and B are obtained by placing/relaxing constraints on the values that the ω parameters can take. Note that all three ω parameters are unconstrained in Model B

on its ω parameters; thus, under Model B episodic selection intensity (ω_{FG}) need not involve positive selection (Figure S2.2C).

Clade-site models: The clade-site models include a site class to model long-term shifts in the intensity of selection pressure. For this class of sites, the phylogeny is divided into two subtrees, and each subtree has an independent ω parameter. Thus, the intensity of section has "shifted" between these subtrees. As we are primarily interested in one subtree, we refer to that one as the "foreground" subtree, and the other as the "background" subtree. Thus the independent ω parameters for subtrees are denoted ω_{FG} and ω_{BG} (Figure S2.3A). The fraction of sites within a gene that have independent ω parameters for subtrees is given by parameter p_{SHIFT} (shown as the proportion, p_2, of site class 2 in Figure 2.3B). Note that the subtree ω parameters (ω_{FG} and ω_{BG}) are independent of the parameters for the sites having a constant selective regime (ω_0 and ω_1); this contrasts with the branch-site models, where the ω_0 and ω_1 are shared between the constant and episodic regimes.

Models C and D have the same basic structure (Figure S2.3A, B), differing only in the restrictions placed on the values of their ω parameters (Figure S2.3C). Model C is obtained by restricting the site-classes of the constant regime such that one of them must have purifying selection ($\omega_0 < 1$) and the other must have neutral evolution ($\omega_1 = 1$). Model D has no restriction on any ω parameters (Figure S2.3C).

4. Likelihood ratio tests. The likelihood ratio test (LRT) is a general method for comparing the fit of two competing hypotheses to a sample of data. The hypotheses must be specified explicitly in the form of a pair of nested models. Specifically, the null hypothesis (H_0) must be a restricted version (special case) of the alternative hypothesis (H_A). The LRT is carried out by comparing twice the difference in log-likelihood scores to a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters between the two models. Note that the LRT only evaluates the difference between a given pair of models; any inadequacies shared by the two models will remain untested.

In this study we employed LRTs to test two generalized classes of evolutionary hypotheses. The first was that sites within NRs were subject to a short-term (episodic) shift in the intensity of natural selection pressure (Figure S2.2). The second was that sites with NRs experienced a long-term shift in the intensity of selection pressure (Figure S2.3). Within each general class of hypothesis, we employed two distinct LRTs; this yields a total of four different LRTs (see Figure S2.4 below).

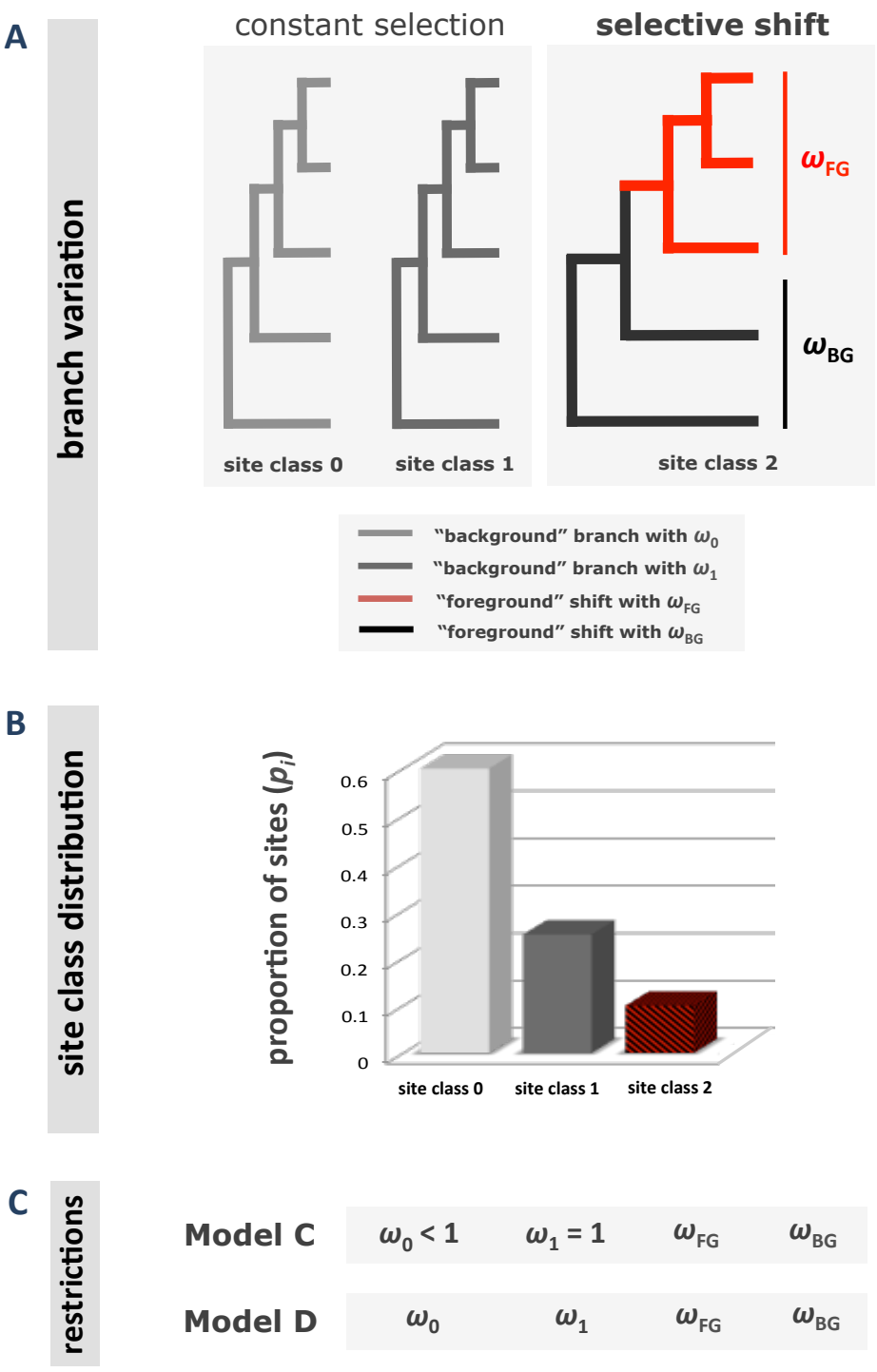

Figure S2.3. Graphical representation of how the ω parameters of clade-site codon models differ among branches and among site classes. (**A**) Two site classes (site classes 0 and 1) specify constant selection pressure over the tree, and a third (site class 2) specifies a long-term selective shift. (**B**) The unconstrained discrete distribution for the three site classes of the codon model. (**C**) Models C and D are obtained by placing/relaxing constraints on the values that the ω parameters can take.

TEST 1 (episodic selection): This LRT is based on two forms of branch-site Models A. In this LRT, the null model constrains the FG branch to neutral evolution (Model A, with $\omega_{FG} = 1$) and the alternative model permits positive selection in the foreground branch (Model A, with $\omega_{FG} > 1$). Thus TEST 1 is intended as a formal test for episodic positive selection along the FG branch. A significant result indicates an episode of positive selection at a fraction of sites within the gene (Zhang et al. 2005; Yang and Dos Reis 2011).

TEST 2 (episodic selection): This LRT compares sites-model M3 (K=2) to branch-site Model B. M3 permits selection pressure to vary among sites, but not among branches. Model B can be viewed as extending M3 by adding a fraction of sites (p_{FG}) that permit selection pressure to change to a unique level in the foreground branch (ω_{FG}). Thus, this is a test for any episodic change in selection (*i.e.,* the episode need not involve $\omega_{FG} > 1$). Formally, a significant result indicates a fraction of sites within a gene (*i.e.*, non-zero p_{FG}) having unique selection along the FG branch (Yang and Nielsen 2002; Zhang et al. 2005).

TEST 3 (selective shift): This LRT compares a relaxed version of sites-model M2a to clade-site Model C (Weadick and Chang, 2012). The null model of this LRT (M2a_rel) serves this function because it permits selection pressure to vary among sites, but not among subtrees. It is "relaxed" because it does not constrain ω_2 to have positive selection; rather, it simply sets $\omega_2 = \omega_{FG} = \omega_{BG}$ for site-class 2. Clade-site Model C is the alternative model because it permits ω to differ between the subtrees at a fraction of sites (p_{SHIFT}). Thus, this pair of models differ in the relationship between ω_{FG} and ω_{BG}. At a subset of sites $\omega_{FG} = \omega_{BG}$ in M2A_rel, whereas $\omega_{FG} \neq \omega_{BG}$ in Model C. TEST 3 is most often used to test for sites having positive selection across an entire clade (*i.e.*, $p_{SHIFT}>0$ and $\omega_{FG}>1$).

TEST 4 (selective shift): This LRT compares sites-model M3 (K=2) to branch-site Model D. M3 permits selection pressure to vary among sites, but not among subtrees. Model D can be viewed as extending M3 by adding a fraction of sites (p_{SHIFT}) that permits selection pressure to change between subtrees. Thus, this is a test for any long-term shift in selection (*i.e.,* the episode need not involve $\omega_{FG} > 1$). A significant result indicates that a fraction of sites within the gene (*i.e.*, $p_{SHIFT} >0$) has experienced a long-term shift in the intensity of selection between subtrees (Bielawski and Yang. 2004).

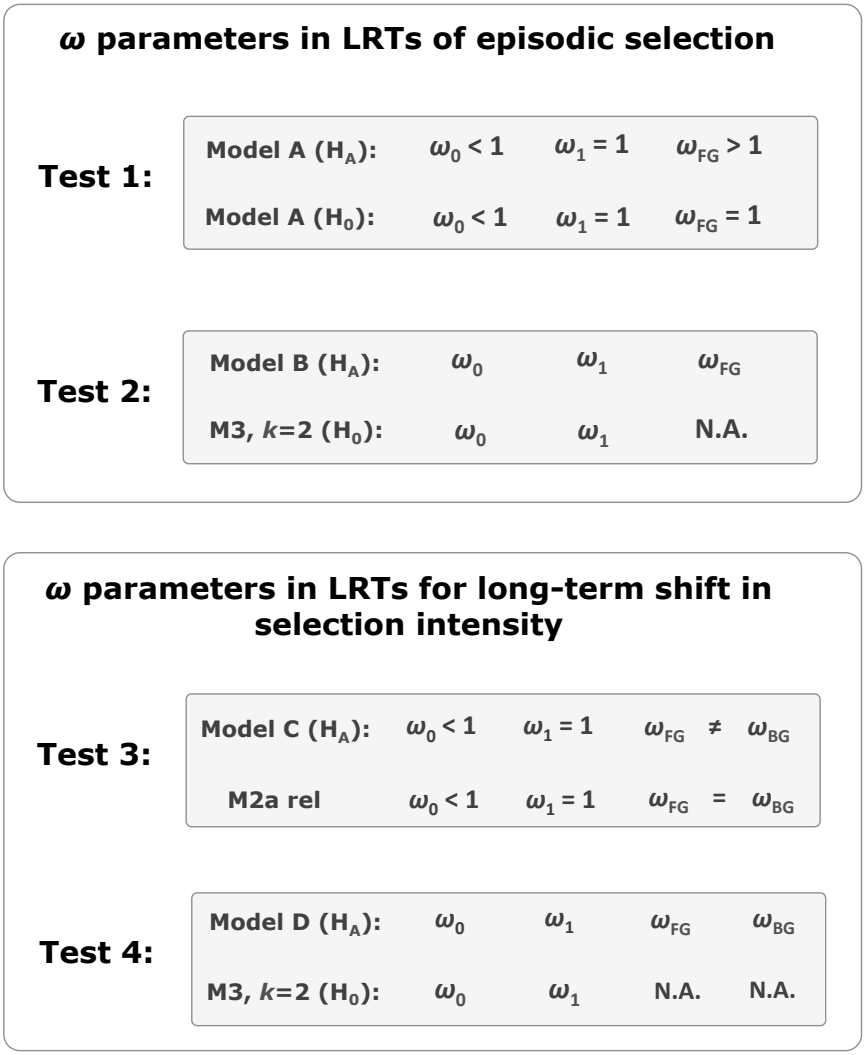

Figure S2.4. The four LRTs are presented in terms of the ω parameters of the involved models. Tests 1 and 2 are for inference of episodic selection (top panel), and Tests 3 and 4 are for inference of long-term selective shifts (bottom panel).

Testing hypotheses in primate NRs: TEST 1 and TEST 2 were used to test each of the three episodic hypotheses (H_1, H_2 & H_3) shown in Figure 1 of the main text. In each case the unique branch associated with each hypothesis was assigned to ω_{FG} in branch-site Models A and B. TEST 3 and TEST 4 were used to test the two hypotheses (H_4 & H_5) about selective shifts (Figure 1 in main text). The primate clades shown in red within Figure 1 of the main text were assigned ω_{FG}, and the remaining branches to ω_{BG}, in Models C and D. The table below summarizes how the hypotheses in Figure 1 in the main text are related to the four tests described above.

Table S2.1. Relationship between the LRTs used in this study and five *a priori* hypotheses about temporal changes in selection presser within primate NRs (H_1 through H_5).

Episodic Selection			Selective Shift	
H_1	H_2	H_3	H_4	H_5
TEST 1	TEST 1	TEST 1	-	-
TEST 2	TEST 2	TEST 2	-	-
-	-	-	TEST 3	TEST 3
-	-	-	TEST 4	TEST 4

5. References.

Bielawski J., Yang Z., 2004 A Maximum Likelihood Method for Detecting Functional Divergence at Individual Codon Sites, with Application to Gene Family Evolution. J. Mol. Evol. 59.

Bielawski, J.P. and Yang, Z., 2005. Maximum likelihood methods for detecting adaptive protein evolution. In *Statistical methods in molecular evolution* (pp. 103-124). Springer New York.

Goldman N., Yang Z., 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11: 725–736.

Nielsen, R. and Yang, Z., 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics, 148(3), pp.929-936.

Muse S. V., Gaut B. S., 1994 A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11: 715–724.

Weadick C. J., Chang B. S. W., 2012 An Improved Likelihood Ratio Test for Detecting Site-Specific Functional Divergence among Clades of Protein-Coding Genes. Mol. Biol. Evol. 29: 1297–1300.

Yang, Z. and Dos Reis, M., 2011. Statistical properties of the branch-site test of positive selection. *Molecular biology and evolution*, *28*(3), pp.1217-1228.

Yang Z., Nielsen R., 2002 Codon-Substitution Models for Detecting Molecular Adaptation at Individual Sites Along Specific Lineages. Mol. Biol. Evol. **19**: 908–917.

Yang, Z., Nielsen, R., Goldman, N. and Pedersen, A.M.K., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, *155*(1), pp.431-449.

Zhang J., Nielsen R., Yang Z., 2005 Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. **22**: 2472–2479.

File S3: Inference of the *NR2C1* amino acid sequence in the ancestor of humans and chimpanzees

1. Overview of the inference method

The ancestral *NR2C1* sequence was inferred from both the DNA data and the amino acid data. The ancestral state was reconstructed for each site of *NR2C1* at the node representing the ancestor of human and chimpanzees. The ancestral state for each variable site was taken as the marginal reconstruction with the highest empirical Bayes posterior probability as implemented in PAML (Yang 2007). Reconstructions for DNA data were based on the GTR model (Yang 1994*a*) with a discrete Gamma (dΓ) correction for among site rate variation (Yang 1994*b*). For the amino acid reconstruction, the data were translated and the calculations were carried out under the WAG exchangeability matrix (Whelan and Goldman, 2001) combined with empirical amino acid frequencies (+F) and the dΓ correction for among site rate variation. With the exception of the state at a single site, the reconstructions were identical. Posterior probabilities were higher under the amino acid model, so the amino acid based reconstruction was chosen as the basis for gene synthesis.

2. The inferred ancestral sequence and the posterior probabilities of the reconstructed amino acid states

The marginal reconstructions of the ancestral states at variable sites under the WAG+F+ dΓ model for amino acid evolution are given below. In addition to each state in the ancestral node, we also provide (*i*) the posterior probability of the ancestral state and (*ii*) the state in the extant sequence for both human and chimpanzee. The full amino acid sequence for the ancestor is provided below the summary of the reconstructed states.

Ancestor → Human

```
Site 242: V 1.000 → M
Site 254: T 0.998 → A *
Site 274: S 1.000 → N
Site 514: G 1.000 → S
```

* The inferred ancestral amino acid at site 254 is different when the sequence is reconstructed under a nucleotide model and translated (see section 3 below).

Ancestor → Chimpanzee

```
37 H 1.000 → R
38 N 1.000 → T
```

```
 94 L 1.000 → M
148 S 1.000 → A
225 T 1.000 → A
```

Complete *NR2C1* amino acid sequence for ancestor:

```
MATIEEIAHQ IIEQQMGEIV TEQQTGQKIQ IVTALDHNTQ GKQFILTNHD
GSTPSKVILA RQDSTPGKVF LTTPDAAGVN QLFFTTPDLS AQHLQLLTDN
SPDQGPNKVF DLCVVCGDKA SGRHYGAVTC EGCKGFFKRS IRKNLVYSCR
GSKDCIINKH HRNRCQYCRL QRCIAFGMKQ DSVQCERKPI EVSREKSSNC
AASTEKIYIR KDLRSPLTAT PTFVTDSEST RSTGLLDSGM FVNIHPSGVK
TESTVLMTSD KAESCQGDLS TLASVVTSLA NLGKTKDLSQ NSNEMSMIES
LSNDDTSLCE FQEMQTNGDV SRAFDTLAKA LNPGESTACQ SSVAGMEGSV
HLITGDSSIN YTEKEGPLLS DSHVAFRLTM PSPMPEYLNV HYIGESASRL
LFLSMHWALS IPSFQALGQE NSISLVKAYW NELFTLGLAQ CWQVMNVATI
LATFVNCLHN SLQQDKMSTE RRKLLMEHIF KLQEFCNSMV KLCIDGYEYA
YLKAIVLFSP DHPGLENMEQ IEKFQEKAYV EFQDYITKTY PDDTYRLSRL
LLRLPALRLM NATITEELFF KGLIGNIRID SVIPHILKME PADYNSQISI
```

3. The inferred ancestral sequence and the posterior probabilities of the reconstructed nucleotide states

The marginal reconstructions of the ancestral states at variable sites under the GTR+ dΓ model for nucleotide evolution are given below. In addition to each state in the ancestral node, we also provide (*i*) the posterior probability of the ancestral state and (*ii*) the state in the extant sequence for both human and chimpanzee. Note that this summery involves more sites than the amino acid summary given above (section 2) because the DNA data contain variation at both synonymous and non-synonymous sites. The full DNA sequence for the ancestor is provided below the summary of the reconstructed states.

Ancestor → Human

```
 724 G 1.000 -> A  (AA: 242 V -> M)
 821 G 1.000 -> A  (AA: 274 S -> N)
1540 G 1.000 -> A  (AA: 514 G -> S)
```

Ancestor → Chimpanzee

```
  30 A 1.000 -> G
 110 A 1.000 -> G  (AA: 37 N -> T)
 113 A 1.000 -> C  (AA: 38 N -> T)
 141 A 1.000 -> G
 280 C 1.000 -> A  (AA: 94 L -> M
 442 T 1.000 -> G  (AA: 148 S -> A)
 603 C 1.000 -> T
 673 A 1.000 -> G  (AA: 225 T -> A)
```

```
760 G 0.848 -> A  (AA: 254 A -> T)*
991 T 1.000 -> C
```

* Nucleotide site 760 has a non-synonymous nucleotide reconstruction that implies a different ancestral amino acid than was reconstructed under the amino acid model described above (section 2).

Complete *NR2C1* nucleotide sequence for ancestor:

```
ATGGCAACCA TAGAAGAAAT TGCACATCAA ATTATTGAAC AACAGATGGG
AGAGATTGTT ACAGAGCAGC AAACTGGGCA GAAAATCCAG ATTGTGACAG
CACTTGATCA TAATACCCAA GGCAAGCAGT TCATTCTGAC AAATCACGAC
GGCTCTACTC CAAGCAAAGT CATTCTGGCC AGGCAAGATT CCACTCCGGG
AAAAGTTTTC CTTACAACTC CAGATGCAGC AGGTGTCAAC CAGTTATTTT
TTACCACTCC TGATCTGTCT GCACAACACC TGCAGCTCCT AACAGATAAT
TCTCCAGACC AAGGACCAAA TAAGGTTTTT GATCTTTGCG TAGTATGTGG
AGACAAAGCA TCAGGACGTC ATTATGGAGC AGTAACTTGT GAAGGCTGCA
AAGGATTTTT TAAAAGAAGC ATCCGAAAAA ATTTAGTATA TTCATGTCGA
GGATCAAAGG ATTGTATTAT TAATAAGCAC CACCGAAACC GCTGTCAATA
CTGCAGGTTA CAGAGATGTA TTGCGTTTGG AATGAAGCAA GACTCTGTCC
AATGTGAAAG AAAACCCATT GAAGTATCAC GAGAAAAATC TTCCAACTGT
GCCGCTTCAA CAGAAAAAAT CTATATCCGA AAGGACCTTC GTAGCCCATT
AACTGCAACT CCAACTTTTG TAACAGATAG TGAAAGTACA AGGTCAACAG
GACTGTTAGA TTCAGGAATG TTCGTGAATA TTCATCCATC TGGAGTAAAA
ACTGAGTCAG CTGTGCTGAT GACATCAGAT AAGGCTGAAT CATGTCAGGG
AGATTTAAGT ACATTGGCCA GTGTGGTTAC ATCATTAGCG AATCTTGGAA
AAACTAAAGA TCTTTCTCAA AATAGTAATG AAATGTCTAT GATTGAAAGC
TTAAGCAATG ATGATACCTC TTTGTGTGAA TTTCAAGAAA TGCAGACCAA
CGGTGATGTT TCAAGGGCAT TTGACACTCT TGCAAAAGCA TTGAATCCTG
GAGAGAGCAC AGCCTGCCAG AGCTCAGTAG CGGGCATGGA AGGAAGTGTA
CACCTAATCA CTGGAGATTC AAGCATAAAT TACACCGAAA AAGAGGGGCC
ACTTCTCAGC GATTCACATG TAGCTTTCAG GCTCACCATG CCTTCTCCTA
TGCCTGAGTA CCTGAATGTG CACTACATTG GGGAGTCTGC CTCCAGACTG
CTGTTCTTAT CAATGCACTG GGCACTTTCG ATTCCTTCTT TCCAGGCTCT
AGGGCAAGAA AACAGCATAT CACTGGTGAA AGCTTACTGG AATGAACTTT
TTACTCTTGG TCTTGCCCAG TGCTGGCAAG TGATGAATGT AGCAACTATA
TTAGCAACAT TTGTCAATTG TCTTCACAAT AGTCTTCAAC AAGATAAAAT
GTCAACAGAA AGAAGAAAAT TATTGATGGA GCACATCTTC AAACTACAGG
AGTTTTGTAA CAGCATGGTT AAACTCTGCA TTGATGGATA CGAATATGCC
TACCTGAAGG CAATAGTACT CTTCAGTCCA GATCATCCAG GCCTAGAAAA
CATGGAACAG ATAGAGAAAT TTCAGGAAAA GGCTTATGTG GAATTCCAAG
ATTATATAAC CAAAACATAT CCAGATGACA CCTACAGGTT ATCCAGACTA
CTACTCAGAT TGCCAGCTTT AAGACTGATG AATGCTACCA TCACTGAAGA
ATTGTTTTTC AAAGGTCTCA TTGGCAATAT ACGAATTGAC AGTGTTATCC
CACATATTTT GAAAATGGAG CCTGCAGATT ATAACTCTCA AATAATAGCA
TT
```

Translation of complete *NR2C1* nucleotide sequence for ancestor:

```
MATIEEIAHQ IIEQQMGEIV TEQQTGQKIQ IVTALDHNTQ GKQFILTNHD
GSTPSKVILA RQDSTPGKVF LTTPDAAGVN QLFFTTPDLS AQHLQLLTDN
SPDQGPNKVF DLCVVCGDKA SGRHYGAVTC EGCKGFFKRS IRKNLVYSCR
GSKDCIINKH HRNRCQYCRL QRCIAFGMKQ DSVQCERKPI EVSREKSSNC
AASTEKIYIR KDLRSPLTAT PTFVTDSEST RSTGLLDSGM FVNIHPSGVK
TESAVLMTSD KAESCQGDLS TLASVVTSLA NLGKTKDLSQ NSNEMSMIES
LSNDDTSLCE FQEMQTNGDV SRAFDTLAKA LNPGESTACQ SSVAGMEGSV
HLITGDSSIN YTEKEGPLLS DSHVAFRLTM PSPMPEYLNV HYIGESASRL
LFLSMHWALS IPSFQALGQE NSISLVKAYW NELFTLGLAQ CWQVMNVATI
LATFVNCLHN SLQQDKMSTE RRKLLMEHIF KLQEFCNSMV KLCIDGYEYA
YLKAIVLFSP DHPGLENMEQ IEKFQEKAYV EFQDYITKTY PDDTYRLSRL
LLRLPALRLM NATITEELFF KGLIGNIRID SVIPHILKME PADYNSQISI
```

References

Whelan, S., Goldman, N., (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18:691–699.

Yang, Z. (1994*a*). Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* 39:105-111.

Yang, Z. (1994*b*). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39:306-314.

Yang, Z., (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol*. 24:1586-1591.

File S4: Sequence validation of human and chimpanzee *NR2C1* cDNA sequences

The sequencing of nonhuman primates has generally relied on deep sequencing of genomic DNA from a small cohort of individuals, without further confirmation by sequencing individual transcripts. We sequenced *NR2C1* transcripts from multiple primates. The four overlapping fragments of chimpanzee *NR2C1* included the sites for all of the amino acid changes implied by the ancestral sequence. None of the amino acid substitutions represented in the chimpanzee or inferred ancestral sequence were identified as nonsynonymous single nucleotide polymorphisms (SNPs). The modest size of our cohort of chimpanzee transcripts means that we cannot exclude the possibility of unobserved polymorphisms in *NR2C1*, but, the concurrence of the primary sequence in at least nine individuals (the three individuals sequenced here, plus the five individuals included in the genome sequencing project) indicates that any nonsynonymous SNPs are likely to be rare. These data also established that a three base deletion leading to the loss of a serine (S) in the human transcript was not present in the chimpanzee, or in the four other primate species we sequenced (**Figure S4.1** below). By querying the dbSNP database (Sherry et al. 2001) we were able to confirm that the human specific amino acid substitutions were not polymorphisms (date of query, March 2013).

References:

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. Oxford University Press; 2001;29: 308–311.

Maynard T, Gopalakrishna D, Meechan D, Paronett E, Newbern J, LaMantia, A. 2013. 22q11 Gene dosage establishes an adaptive range for sonic hedgehog and retinoic acid signaling during early development. Hum Mol Genet 22(2):300-312.

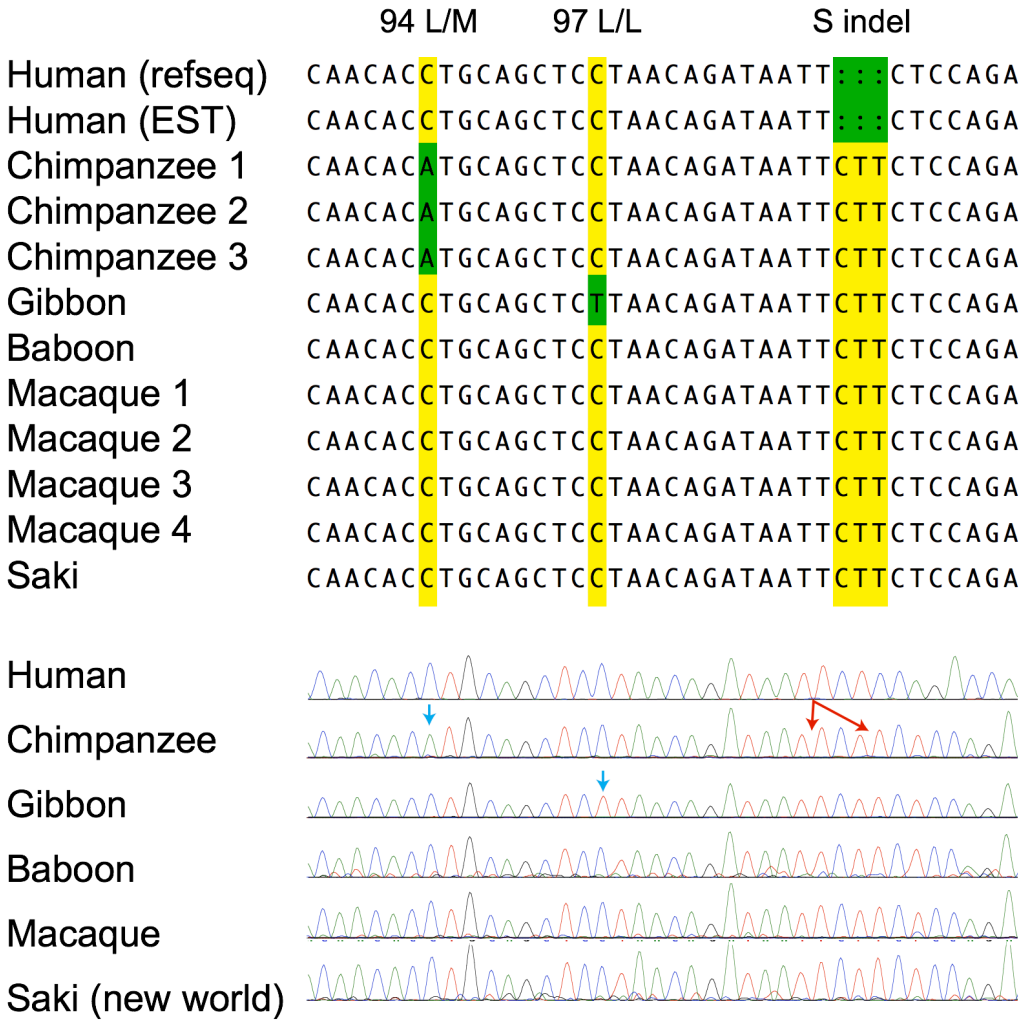

Figure S4.1. Sequencing of Nr2c1 transcript cDNAa from primate cortex. Examples of single nucleotide polymorphisms are identified in chromatograms by red arrows; the human specific deletion in the N-terminal domain is marked by red arrows.

File S5: Testing for recombination within NRs by using the GARD-MBP method.

Because recombination can negatively impact LRTs (Anisimova et al. 2003), we tested for a history of recombination in each of the nine genes that we identified as associated with a shift in selection intensity at either the origin of great apes (H_4) or the origin of the human-chimpanzee clade (H_5). We used the GARD-MBP method (Kosakovsky Pond et al. 2006), which Bay and Bielawski (2011) found was powerful under evolutionary scenarios relevant to functional divergence. GARD employs a genetic algorithm (GA) to search for and identify potential within-gene recombination breakpoints according to phylogenetic incongruence. Overall significance for a history of within-gene homologues recombination is based on the principle of an "omnibus test"; an alignment is tested twice, by passing over it from both directions, and a gene is considered significant only if both tests are significant (Kosakovsky Pond et al. 2006). We employed the GA to search for the multi-breakpoints (MBP) under a full GTR model (Yang 1994*a*) for nucleotide evolution. The method was carried out as implemented in the Datamonkey server (Delport et al. 2010: www.Datamonkey.org). The *p*-values provided by the Datamonkey server are corrected for multiple testing. Results for each of the nine candidate NR genes are presented in **Table S3.1** below.

Table S3.1. Results of testing for within gene homologous recombination by using the GARD-MBP method

	inferred	*p*-value		
Gene Name	**breakpoint(s)**	**LHS**	**RHS**	**significance (0.05)**
ESRRA	974	0.4112	0.02080	n.s.
ESRRB	647	1.0	0.00042	n.s.
NR1D1	none	na	na	na
NR2C1	none	na	na	na
NR2E3	818	**0.032**	**0.0036**	**
	899	0.2672	0.8952	n.s.
PGR	245	0.1236	0.4086	n.s.
	1075	1.0	0.0006	n.s.
	1423	0.0006	0.0642	n.s.
PPARG	446	0.0002	0.4052	n.s.
RARG	548	0.0002	0.5308	n.s.
RORA	165	0.0002	0.7574	n.s.

Note: LHS denotes tests carried out from "left-hand side". RHS denotes tests carried out from "right-hand side". Significance under the GARD-MB method requires omnibus statistical significance; *i.e.*, both LHS and RHS must be significant. *p*-values produced by the GARD-MBP method include corrections for applying multiple tests to the same gene.

References

Anisimova, M., Nielsen, R., Yang, Z., (2003). Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 164:1229–1236.

Bay, R.A., Bielawski, J.P., (2011). Recombination detection under evolutionary scenarios relevant to functional divergence. *J. Mol. Evol.* 73:273–286.

Delport, W., Poon, A. F., Frost, S. D., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, *26*: 2455-2457.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B, Woelk, C.H., and Frost, S.D.W. (2006). Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm. *Mol. Biol. Evol.* 23:1891-1901

Yang, Z. (1994*a*). Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* 39:105-111.

File S6: Assessing the impact of among-site variation in the baseline rate of DNA/RNA substitution on the estimation of the ω parameter.

The rate of substitution can vary among sites according to processes that are unrelated to the intensity of natural selection acting at the amino acid level of evolution. The major processes are purifying selection acting at the DNA or RNA level of evolution, or recombination that gives rise to difference among sites in the underlying tree topology and branch lengths. Both processes can lead to among-site differences in the synonymous rate, and estimates of amino acid level selection (via the ω parameter) using models that do not permit such variability can be biased when the synonymous rate variation is large enough (Kosakovsky Pond and Muse 2005; Scheffler et al. 2006; Rubenstein et al. 2011).

Because we detected variation in the branch length scale parameters for different structural domains (main text and File S1) we investigated the potential impact of such variation on estimates of ω for NR genes. We employed the multi-layer codon model of Rubenstein et al. (2011), as it permits any source of variation to act at either the DNA or RNA level; this is a more flexible framework than a model which restricts this process to just the synonymous opportunities between codons permitted by the genetic code. To assess the potential impact of baseline variability in the DNA/RNA substitution rate, we estimated the posterior mean value of ω for each site under two versions of codon model M8; one version was the standard M8 (no variation in the baseline DNA/RNA substitution rate) and the other was a multi-layer M8 with baseline DNA/RNA substitution rates modeled according to a discrete gamma distribution. Posterior mean ω's for each site were compared by plotting them against each other in a scatter plot. When the DNA/RNA substitution rate does not impact estimate of ω, the data should fall along the diagonal.

We compared the posterior mean ω's in each of the nine genes that we identified as associated with a shift in selection intensity at either the origin of great apes (H_4) or the origin of the human-chimpanzee clade (H_5). We found that in seven of the nine genes (*ESSRA, ESSRB, NR2C1, NR2E3, PRG, RARG* and *RORA*) the vast majority of estimates for ω (and in many cases all of them) fell along the diagonal. This indicated little to no impact on ω. This was not the case, however, for *NR1D1* and *PPARG*. The plot for *NR1D1* indicated that estimates under the multi-layer model tended to be a little lower than under the standard codon model. The plot for *PPARG* indicated a much larger systematic difference, with the multi-layer model consistently yielding substantially lower estimates than the standard codon model. All nine plots are presented in **Figure S6.1** below.

References

Kosakovsky Pond S.L., Muse S.V., (2005) Site-to-site variation of synonymous substitution rates. *Mol. Biol. Evol.* 22:2375-2385.

Rubinstein N.D., Doron-Faigenboim A., Mayrose I., Pupko T., (2011) Evolutionary models accounting for layers of selection in protein-coding genes and their impact on the inference of positive selection. *Mol. Biol. Evol*. 28:3297-308.

Scheffler K., Martin D.P., Seoighe C. (2006). Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 2006; 22:2493-2499.

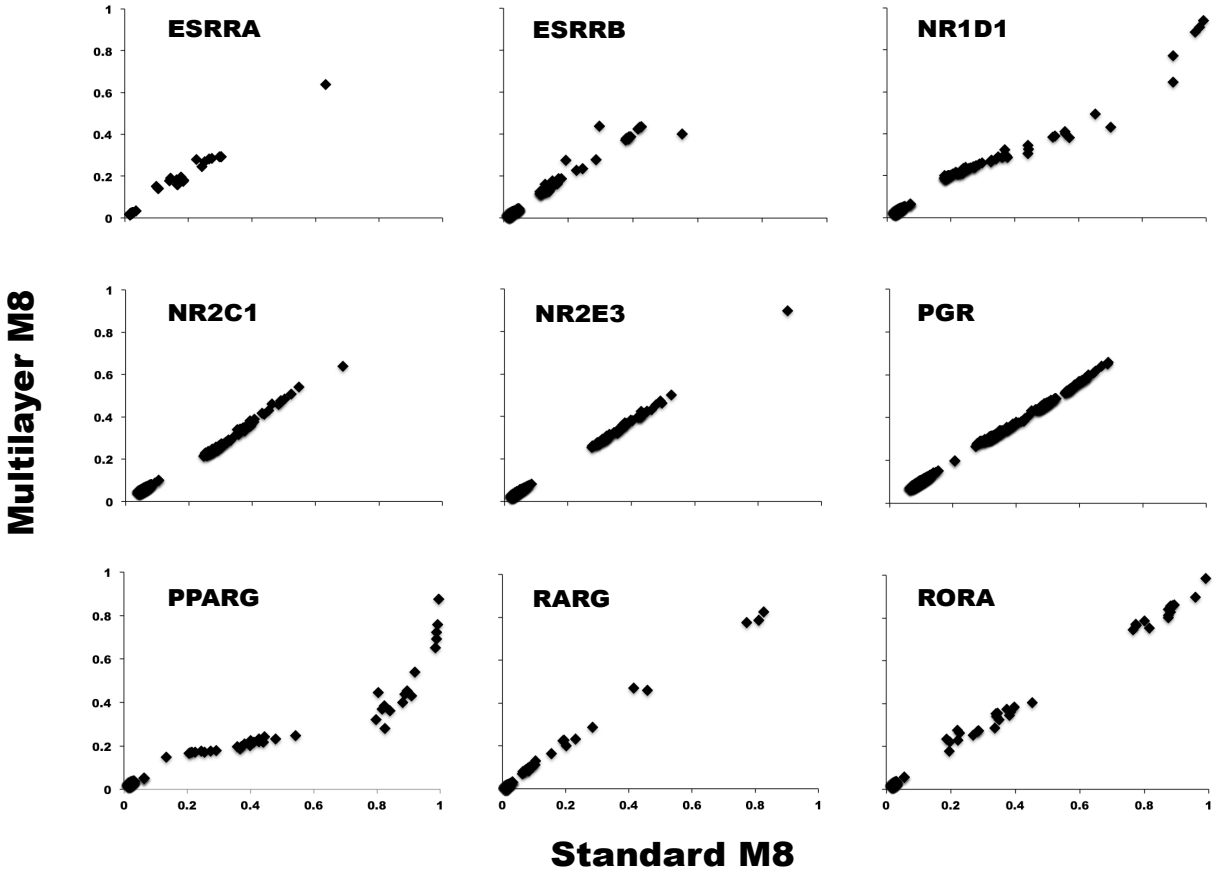

Figure S6.1. Comparison of posterior mean ω's under the standard M8 to those inferred under a multi-layer M8 model in each of nine candidate NRs. The multi-layer codon model of Rubenstein et al. (2011) permits among-site variability in the baseline rate of DNA/RNA substitution, whereas the standard M8 does not. Estimates of ω under the standard M8 model can be negatively impacted when there is substantial variation in in the baseline rate of DNA/RNA substitution. To assess the sensitivity of ω to model assumptions about among-site variability rate of DNA/RNA substitution we plotted the posterior mean ω's against each other for these two models. When ω is not dependent on assumptions about the baseline rate of DNA/RNA substitution, the estimates should fall along the diagonal. We observed this in seven of the genes (*ESRRA, ESRRB, NR2C1, NR2E3, PGR, RARG* and *RORA*). However, *NRD21* showed some minor sensitivity to the model, and estimates of ω for *PPARG* were highly sensitive to model assumptions. On a site-wise basis, there was a systematic overestimation of ω under the standard M8 model, although in no case was ω>1.

File S7: Ranking the top four candidates for AGR-based experimental assays.

We identified four genes (*ESRRB, PGR, RORA* and *NR2C1*) that (*i*) had an alteration in selection pressure associated with the origin of the great ape and human-chimpanzee clades, and (*ii*) passed a series of reliability and robustness analyses. AGR-based characterization of each of these genes was not within the scope of this study. Furthermore, each one was not an equally promising candidate. Hence, we ranked them based on our assessment of the signal for a change in selection pressure, the extent to which the biological role of each gene has been characterized, and the capacity to carry out *in vitro* assays for functional effects of amino acid substitutions.

We begin with *ESRRB*, which we ranked fourth mainly because we were looking for a gene whose pattern of sequence change suggested there could be functional divergence in the human lineage. This gene appears to have a relatively large fraction of sites (~28%) that experienced a shift in selection intensity, in the great ape clade. *ESRRB* is interesting, however, as it is a known regulator of pluripotency (Festuccia et al., 2012; Martello et al., 2012).This gene also appears to be important in maintaining a hypoxic environment, which is essential for tumor cell growth and persistence, by interacting with and inducing hypoxia-inducible factor (Ao et al. 2008), making *ESRRB* an attractive target for cancer therapy (Cheung et al. 2005; Yu et al. 2008; Sengupta et al. 2014). However, we considered it the least promising candidate because we were unable to obtain consistent signal for a selective shift at the origin of Humans and Chimps (i.e., LRT-3 was significant for H_5, but LRT-4 was not).

We ranked the progesterone receptor gene (*PGR*) third. Progesterone plays a central role in the biology of childbirth (Csapo 1961; Graham and Clarke 1997; Henson 1998) and progesterone-receptor interaction is crucial for establishing and maintaining a pregnancy (Challis et al. 2000). The effects of progesterone are controlled by interactions between *PGR* and its coregulators (Lonard and O'Malley 2007). This cofactor interaction is crucial for the initiation of labor (Condon et al. 2003), so it is noteworthy that most substitutions in the human *PGR* gene variant occurred in the N-terminal domain of *PGR* where cofactor binding takes place (Chen et al. 2008; this study). Chen et al. (2008) hypothesized that *PGR* plays a role in primate reproductive

diversity, and that amino acid changes in its cofactor interaction surfaces have been subject to positive Darwinian selection. This gene was not rated higher because it is the subject of a preexisting investigation.

We ranked *RORA* second. This gene offers protection to neurons and glial cells from the degenerative effects of oxidative stress (Boukhtouche et al. 2006; Jolly et al. 2011) and is widely expressed in a number of brain regions including the cerebral cortex, hypothalamus, and thalamus (Ino 2004). Recent papers have implicated *RORA* in a variety of clinical conditions, including attention-deficit hyperactivity disorder (Neale et al. 2008), bipolar disorder (Le-Niculescu et al. 2009), depression (Garriock et al. 2010; Terracciano et al. 2010), autism (Sarachana et al. 2011) and and a GWAS identified a significant risk locus in this gene for posttraumatic stress disorder (Logue et al. 2013; Amstadter et al. 2013). Because of this evidence, plus the strong signal for positive Darwinian selection within the great ape clade, we ranked *RORA* above *PGR* and *ESSRB* as a candidate gene.

Our top candidate, *NR2C1,* belongs to a subtype of NRs known as orphan receptors for which the endogenous ligand (if any) has yet to be identified (Lee and Chang 1995). Originally named the testicular receptor 2 (*TR2*) because it was first isolated from modern human testis and prostate (Chang and Kokontis 1988; Anderson et al. 2012), its expression in embryonic stem cells and in pluripotent cell culture lines indicates it plays a role in early embryonic development (Hu et al. 2002). It is one of a handful of genes implicated in the regulation of the pluripotentiality of stem cell populations in the embryo, and in neural stem cells in particular (Lee and Chang 1995; Hu et al. 2002; Lee et al. 2002; Shyr et al. 2009). In addition, *NR2C1* has been shown to regulate the expression of *Oct4* and *Nanog*, two transcription factors essential for maintaining the pluripotentiality of embryonic stem cells (Pikarsky et al. 1994; Niwa et al. 2000; Boiani et al. 2002). Given its role in maintaining the pluripotentiality of stem cells and in neural differentiation, and a signal for positive selection within the human-chimpanzee clade, we hypothesized that *NR2C1* has played a role in the evolution of a developmental system(s) relevant to the anatomical and physiological characteristics that distinguish humans and chimpanzees from the other great apes.

In Phase 3 of this study we took the first steps to evaluate the hypothesis that amino

acid substitutions within *NR2C1* since the last common ancestor of humans and chimpanzees have affected its capacity to regulate pluripotentiality.

References

Amstadter AB, Sumner JA, Acierno R, Ruggiero KJ, Koenen KC, Kilpatrick DG, et al. Support for association of RORA variant and post traumatic stress symptoms in a population-based study of hurricane exposed adults. Mol Psychiatry. 2013;18: 1148–1149. doi:10.1038/mp.2012.189

Anderson AM, Carter KW, Anderson D, Wise MJ. Coexpression of Nuclear Receptors and Histone Methylation Modifying Genes in the Testis: Implications for Endocrine Disruptor Modes of Action. Veitia RA, editor. PLoS ONE. 2012;7: e34158. doi:10.1371/journal.pone.0034158

Boiani M. Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes & Development. 2002;16: 1209–1219. doi:10.1101/gad.966002.

Boukhtouche F, Vodjdani G, Jarvis CI, Bakouche J, Staels B, Mallet J, et al. Human retinoic acid receptor-related orphan receptor alpha1 overexpression protects neurones against oxidative stress-induced apoptosis. J Neurochem. Blackwell Science Ltd; 2006;96: 1778–1789. doi:10.1111/j.1471- 4159.2006.03708.x

Challis JRG, Matthews SG, Gibb W, Lye SJ. Endocrine and paracrine regulation of birth at term and preterm. Endocrine Reviews. Endocrine Society; 2000;21: 514–550. doi:10.1210/edrv.21.5.0407

Chang C, Kokontis J. Identification of a new member of the steroid receptor super-family by cloning and sequence analysis. Biochemical and Biophysical Research Communications. 1988;155: 971–977.

Chen C, Opazo JC, Erez O, Uddin M, Santolaya-Forgas J, Goodman M, et al. The human progesterone receptor shows evidence of adaptive evolution associated with its ability to act as a transcription factor. Mol Phylogenet Evol. 2008;47: 637– 649. doi:10.1016/j.ympev.2007.12.026

Cheung CP, Yu S, Wong KB, Chan LW, Lai FMM, Wang X, et al. Expression and Functional Study of Estrogen Receptor-Related Receptors in Human Prostatic Cells and Tissues. The Journal of Clinical Endocrinology & Metabolism. 2005;90: 1830–1844. doi:10.1210/jc.2004-1421

Condon JC, Jeyasuria P, Faust JM, Wilson JW, Mendelson CR. A decline in the levels of progesterone receptor coactivators in the pregnant uterus at term may antagonize progesterone receptor function and contribute to the initiation of parturition. Proc Natl Acad Sci USA. National Acad Sciences; 2003;100: 9518– 9523. doi:10.1073/pnas.1633616100

Csapo AI. The onset of labour. Lancet. 1961;2: 277–280. Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, and Chambers I. 2012. Cell Stem Cell 11, 477–490.

Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, Chambers I. Esrrb is a direct Nanog target gene that can

substitute for Nanog function in pluripotent cells. Cell Stem Cell, 2012;11(4):477–490.
Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genomewide association study of citalopram response in major depressive disorder. Biol Psychiatry. Elsevier; 2010;67: 133–138.
Graham JD, Clarke CL. Physiological action of progesterone in target tissues. Endocrine Reviews. Endocrine Society; 1997;18: 502–519.
Henson MC. Pregnancy maintenance and the regulation of placental progesterone biosynthesis in the baboon. Hum Reprod Update. 1998;4: 389–405.
Hu Y-C, Shyr C-R, Che W, Mu X-M, Kim E, Chang C. Suppression of estrogen receptor-mediated transcription and cell growth by interaction with TR2 orphan receptor. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2002;277: 33571–33579.
Ino H. Immunohistochemical characterization of the orphan nuclear receptor ROR alpha in the mouse nervous system. J Histochem Cytochem. 2004;52: 311–323.
Jolly S, Journiac N, Naudet F, Gautheron V, Mariani J, Vernet-der Garabedian B. Cell-autonomous and non-cell-autonomous neuroprotective functions of RORα in neurons and astrocytes during hypoxia. J Neurosci. Society for Neuroscience; 2011;31: 14314–14323. doi:10.1523/JNEUROSCI.1443-11.2011
Kaya AO, Gunel N, Benekli M, Akyurek N, Buyukberber S, Tatli H, et al. Hypoxia inducible factor-1 alpha and carbonic anhydrase IX overexpression are associated with poor survival in breast cancer patients. J BUON. 2012;17: 663– 668.
Le-Niculescu H, Patel SD, Bhat M, Kuczenski R, Faraone SV, Tsuang MT, et al. Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. Am J Med Genet B Neuropsychiatr Genet. Wiley Subscription Services, Inc., A Wiley Company; 2009;150B: 155–181. doi:10.1002/ajmg.b.30887
Lee HJ, Chang C. Identification of Human TR2 Orphan Receptor Response Element in the Transcriptional Initiation Site of the Simian Virus 40 Major Late Promoter. J Biol Chem. 1995;270: 5434–5440. doi:10.1074/jbc.270.10.5434
Lee Y-F, Lee H-J, Chang C. Recent advances in the TR2 and TR4 orphan receptors of the nuclear receptor superfamily. J Steroid Biochem Mol Biol. 2002;81: 291–308.
Logue MW, Baldwin C, Guffanti G, Melista E, Wolf EJ, Reardon AF, et al. A genome-wide association study of post-traumatic stress disorder identifies the retinoid-related orphan receptor alpha (RORA) gene as a significant risk locus. Mol Psychiatry. 2013;18: 937–942. doi:10.1038/mp.2012.113
Lonard DM, O'Malley BW. Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation. Mol Cell. Elsevier; 2007;27: 691–700. doi:10.1016/j.molcel.2007.08.012
Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Gottgens B, Niwa H, and Smith A. 2012. Cell Stem Cell 11, 491–504.
Neale BM, Lasky-Su J, Anney R, Franke B, Zhou K, Maller JB, et al. Genome- wide association scan of attention deficit hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. Wiley Subscription Services, Inc., A Wiley Company; 2008;147B: 1337–1344. doi:10.1002/ajmg.b.30866

Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, et al. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 2005;3: e170. doi:10.1371/journal.pbio.0030170

Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 2000;24: 372–376. doi:10.1038/74199.

Pikarsky E, Sharir H, Ben-Shushan E, Bergman Y. Retinoic acid represses Oct-3/4 gene expression through several retinoic acid-responsive elements located in the promoter-enhancer region. Mol Cell Biol. 1994;14: 1026–1038.

Sarachana T, Xu M, Wu R-C, Hu VW. Sex hormones in autism: androgens and estrogens differentially and reciprocally regulate RORA, a novel candidate gene for autism. Mattson M, editor. PLoS ONE. 2011;6: e17116. doi:10.1371/journal.pone.0017116

Sengupta D, Bhargava DK, Dixit A, Sahoo BS, Biswas S, Biswas G, et al. ERRβ signalling through FST and BCAS2 inhibits cellular proliferation in breast cancer cells. British Journal of Cancer. 2014;110: 2144–2158. doi:10.1038/bjc.2014.53

Shyr C-R, Kang H-Y, Tsai M-Y, Liu N-C, Ku P-Y, Huang K-E, et al. Roles of Testicular Orphan Nuclear Receptors 2 and 4 in Early Embryonic Development and Embryonic Stem Cells. Endocrinology. 2009;150: 2454–2462. doi:10.1210/en.2008-1165

Terracciano A, Tanaka T, Sutin AR, Sanna S, Deiana B, Lai S, et al. Genome-wide association scan of trait depression. Biol Psychiatry. Elsevier; 2010;68: 811–817. doi:10.1016/j.biopsych.2010.06.030

Yu S, Wong YC, Wang XH, Ling MT, Ng CF, Chen S, et al. Orphan nuclear receptor estrogen-related receptor-β suppresses in vitro and in vivo growth of prostate cancer cells via p21WAF1/CIP1 induction and as a potential therapeutic target in prostate cancer. Oncogene. 2007;27: 3313–3328. doi:10.1038/sj.onc.1210986