LETTER

doi:10.1038/nature10595

Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB

Vincent J. Lynch[1], Gemma May[1] & Günter P. Wagner[1]

There is an emerging consensus that gene regulation evolves through changes in *cis*-regulatory elements[1,2] and transcription factors[3–6]. Although it is clear how nucleotide substitutions in *cis*-regulatory elements affect gene expression, it is not clear how amino-acid substitutions in transcription factors influence gene regulation[4–10]. Here we show that amino-acid changes in the transcription factor CCAAT/enhancer binding protein-β (CEBPB, also known as C/EBP-β) in the stem-lineage of placental mammals changed the way it responds to cyclic AMP/protein kinase A (cAMP/PKA) signalling. By functionally analysing resurrected ancestral proteins, we identify three amino-acid substitutions in an internal regulatory domain of CEBPB that are responsible for the novel function. These amino-acid substitutions reorganize the location of key phosphorylation sites, introducing a new site and removing two ancestral sites, reversing the response of CEBPB to GSK-3β-mediated phosphorylation from repression to activation. We conclude that changing the response of transcription factors to signalling pathways can be an important mechanism of gene regulatory evolution.

A major challenge in evolutionary and developmental biology is discovering the molecular mechanisms of gene regulatory evolution. Although it is clear that changes in transcription factors have contributed to the evolution of gene regulation[7–11], most studies of gene regulatory evolution have focused on *cis*-regulatory elements[1,2,12–15]. Thus, how mutations in transcription factors generate new regulatory functions is not well understood.

The mammalian transcription factor CEBPB plays an essential role during pregnancy[16–18], including physically and functionally interacting with FOXO1A to activate prolactin (PRL) expression in endometrial stromal cells (ESCs)[19–21] (Supplementary Fig. 1a). To test if this function evolved coincident with the origin of pregnancy in placental mammals, we assayed chicken, opossum and human CEBPB (LAP) and FOXO1A for their ability to transactivate luciferase expression from the ESC *PRL* enhancer. We found that each species' CEBPB moderately transactivated luciferase expression in HeLa and ESC (Fig. 1a, b). Co-expressed human FOXO1A and CEBPB cooperatively transactivated luciferase expression approximately 3.6-fold (Fig. 1a, b). In stark contrast, co-expression of CEBPB and FOXO1A from opossum and chicken were unable to transactivate luciferase expression cooperatively (Fig. 1a, b).

These results suggest that the cooperative CEBPB/FOXO1A interaction evolved in placental mammals. To test this hypothesis we used maximum likelihood to reconstruct the sequence of the ancestral therian (AncTherian) and eutherian (AncEutherian) CEBPB and FOXO1A genes, synthesized them using human codon-usage and tested their function in the assay described above. We found that the AncEutherian CEBPB/FOXO1A complex strongly transactivated luciferase expression whereas the AncTherian complex was unable to do so (Fig. 1a, b). This derived cooperativity in placental mammals, however, does not result from a newly evolved physical interaction between CEBPB and FOXO1A because both the AncEutherian and AncTherian CEBPB proteins are able to interact physically with ancestral FOXO1A proteins (Fig. 1c).

Previous studies have characterized several functionally distinct regions in CEBPB including three activation domains (ADM1–3), two internal regulatory domains (RD1–2) and a DNA-binding domain (DBD)[22–24]. To determine which domains and amino-acid substitutions are responsible for the derived function of CEBPB, we made

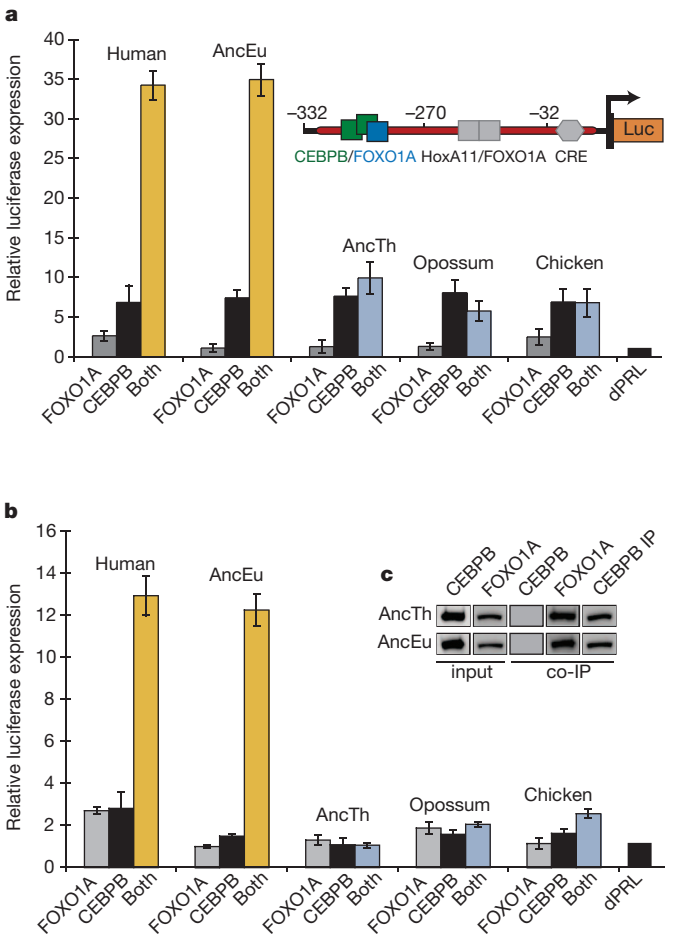

Figure 1 | CEBPB cooperatively regulates gene expression with FOXO1A in placental mammals. CEBPB and FOXO1A from human and the reconstructed ancestor of eutherian mammals (AncEu) cooperatively transactivate luciferase expression from the decidual prolactin promoter (dPRL) when co-transfected into HeLa (**a**) or human ESCs (**b**). Inset, the structure of the luciferase reporter vector and experimentally characterized transcription-factor binding sites. CEBPB and FOXO1A from non-mammals, including the reconstructed ancestor of therian mammals (AncTh), do not show cooperative upregulation. Luciferase values are shown as fold changes (mean ± s.e.m., $n = 6$) relative to the reporter control (dPRL). **c**, Both the ancestral therian and eutherian CEBPB proteins directly interact. CEBPB-V5/His and FOXO1A–Flag were expressed in HeLa cells and immunoprecipitated with anti-Flag agarose beads followed by DNase treatment.

[1]Yale Systems Biology Institute and Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06511, USA.

amino (N)-terminal truncation mutants of the AncTherian and AncEutherian CEPBPB genes and tested their ability to transactivate luciferase expression when co-expressed with FOXO1A. Deletion of ADM1 (Δ39) or ADM1–2 (Δ109) abrogated the transactivation ability of the AncEutherian protein whereas the ADM1–3 (Δ131) deletion restored transactivation. Deletion of RD1 (Δ175) and the region between RD1 and RD2 (Δ200) had little effect on the enhanced activation of Δ131. Deletion of RD2 (Δ236), however, reduced the transactivation ability of the AncEutherian protein to that observed for the AncTherian protein (Fig. 2a), indicating that RD2 plays a critical role in PRL regulation.

To test if the five eutherian-specific amino-acid changes within this region are responsible for the derived function in placental mammals, we reverted the derived amino acids in AncEutherian RD2 back to their ancestral state in the AncTherian protein and tested their ability to transactivate luciferase expression. We found that back mutation of derived sites one and two had no effect on transactivation, whereas back mutation of sites three, four and five resulted in a pronounced loss of transactivation ability (Fig. 2b). Next, we tested whether forward mutation of sites three, four and five was sufficient to impart a transactivation function to the AncTherian CEBPB. Unlike the progressive loss of function observed for reverse mutations, forward mutation of sites three and four, alone or in combination, had no effect on the ability of AncTherian CEBPB to transactivate luciferase expression (Fig. 2b). Forward mutation of site five only modestly enhanced transactivation; however, forward mutation of sites three and four in combination with site five dramatically enhanced transactivation by the AncTherian CEBPB (Fig. 2b). Thus, although the amino-acid substitutions at sites three and four contribute to the derived regulatory ability of CEBPB in placental mammals, they are dependent on the amino-acid substitution at site five.

Previous studies have shown that CEBPB is phosphorylated in response to cell-signalling, differentiation[25] and cAMP/PKA stimulation[26]. We mapped known and predicted phosphorylation sites in CEBPB and identified several differences between non-placental and placental mammals, including the loss of two ancestral serine phosphosites at sites three (S3A) and four (S4A) and the gain of a serine phosphosite at site five (N5S) in placental mammals. These sites are generally conserved at state S-(S/T)-N outside placental mammals and at state A-A-S within placental mammals, suggesting they are functionally constrained (Fig. 2c). Remarkably, the eutherian-specific serine at site five has previously been shown to be phosphorylated by GSK-3β kinase in rat CEBPB[27]. We confirmed that this site was phosphorylated in the AncEutherian CEBPB using nano-LC-ESI-MS/MS on purified protein *in vitro* phosphorylated with MAPK and GSK-3β kinases. Automated analysis of the tandem mass spectrometry (MS/MS) spectra indicated that the derived serine at site five is phosphorylated by GSK-3β kinases.

To test if phosphorylation by GSK-3β plays a role in potentiating transactivation by CEBPB, we repeated our luciferase reporter assays with the ancestral CEBPB and FOXO1A proteins but blocked GSK-3β phosphorylation with the kinase inhibitor 6-bromoindirubin 3′-oxime (BIO); if phosphorylation by GSK-3β is functionally important for transactivation of reporter gene expression by CEBPB, then inhibiting this kinase should block transactivation. Indeed, we found that inhibiting GSK-3β blocked the ability of the AncEutherian CEBPB to transactivate luciferase expression from the dPRL enhancer (Fig. 2b). Unexpectedly, however, inhibiting GSK-3β dramatically potentiated transactivation by AncTherian CEBPB. These results indicate that the AncEutherian CEBPB is activated by GSK-3β, whereas the AncTherian CEBPB is repressed by GSK-3β (Fig. 2b).

To determine which amino acids are responsible for interpreting GSK-3β-mediated phosphorylation as an inhibiting signal in AncTherian CEBPB and an activating signal in AncEutherian CEBPB, we compared the ability of back and forward mutated proteins to cooperatively transactivate luciferase expression with FOXO1A in the presence of BIO. We found that back mutation of site three, four and three/four in the AncEutherian CEBPB protein, which restores ancestral putative

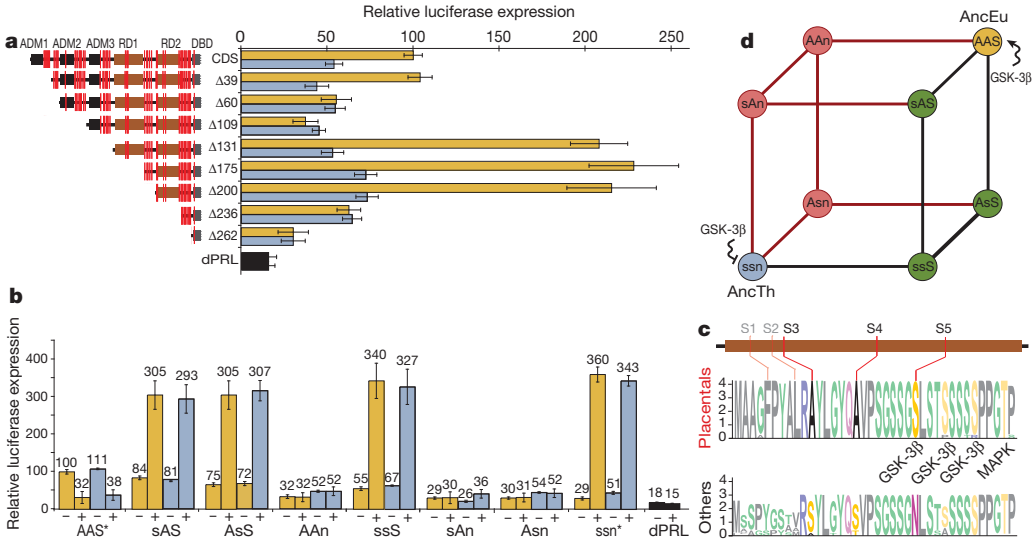

Figure 2 | CEBPB evolved a novel GSK-3B phosphorylation site in an internal regulatory domain. **a**, N-terminal deletion mutants of ancestral eutherian (yellow) and ancestral therian (blue) CEBPB differentially transactivate the dPRL-luciferase reporter when co-transfected with FOXO1A. Deletion of regulatory domain 2 (RD2; Δ236) in the ancestral eutherian CEBPB protein reduces its transactivating ability to that of the ancestral therian protein. Luciferase values are shown as fold changes relative to the full-length (CDS) ancestral eutherian CEBPB protein (mean ± s.e.m., $n = 6$). Red lines indicate the position of eutherian-specific amino-acid changes. **b**, Effect of forward and back mutation of sites three, four, and five in the ancestral therian (blue) and eutherian (yellow) CEBPB on transactivation in the presence of BIO (+) or BIO control (−) media. Luciferase values are shown relative to wild-type ancestral eutherian CEBPB (mean ± s.e.m., $n = 6$); *wild-type proteins; lower-case letters indicate amino-acid state in the ancestral therian whereas upper-case letters indicate the state in the ancestral eutherian. **c**, Amino-acid substitutions at sites three, four and five reposition the location of key phosphosites. RD2 logos are shown for placental mammals and other species with the location of S/T phosphosites highlighted in yellow with the name of the phosphorylating kinases; nano-liquid chromatography–electrospray ionization/tandem mass spectrometry of *in vitro* phoshorylated ancestral eutherian CEBPB confirmed site five (S5) is phosphorylated by GSK-3β. **d**, Mutational pathway between the ancestral therian (AncTh, blue) and eutherian (AncEu, yellow) functions. Each corner of the cube represents a possible intermediate state for amino acids at sites three, four and five; amino-acid identity at each state is shown as single letter amino code. Red edges show unlikely evolutionary paths through intermediate states that cannot transactivate reporter gene expression (red nodes), black edges show paths through functional intermediate states (green nodes).

phosphorylation sites that were lost in the stem lineage of placental mammals, recapitulated the dramatic enhancement of transactivation observed for AncTherian CEBPB in the presence of BIO (Fig. 2b). Conversely, forward mutations of sites three, four and three/four in the AncTherian CEBPB protein, which abolish the ancestral phosphorylation sites but do not introduce the derived phosphosite at site five, blocked its ability to transactivate reporter gene expression regardless of BIO treatment, leading to an essentially functionless protein in this assay (Fig. 2b). Incorporating a forward mutation at site five increased the transactivation ability of AncTherian CEBPB in the absence of BIO while maintaining a strong enhancement of transactivation with BIO treatment; however, transactivation with BIO was lost upon forward mutations of sites three and four (Fig. 2b). We conclude that the N5S substitution probably came before S3A and S4A because these substitutions lead to non-functional proteins when not paired with N5S and evolution is unlikely to pass through such non-functional intermediates (Fig. 2d).

Previous studies suggest that phosphorylation of RD2 induces a conformation change in CEBPB that disrupts intramolecular contacts between RD1–2, ADM1-3 and the bZIP/DNA-binding domain, leading to activation of the intrinsically repressed protein[17,22–24,28]. To infer the structural consequences of repositioning phosphorylation sites, we generated structural models of unphosphorylated and phosphorylated proteins using Rosetta/Robetta[29]. The model of the unphosphorylated AncEutherian protein suggests that ADM1–3 and RD1–2 are intrinsically unstructured and collapsed into a tight knot-like bundle in close contact with the helical bZIP/DNA-binding domain (Fig. 3a) similar to previous molecular and biochemical models[22–24]. In contrast, ADM1–3 and RD1–2 associate with neither each other nor the bZIP/DNA-binding domain in the model of the phosphorylated AncEutherian CEBPB (Fig. 3a), freeing residues critical for nuclear localization and DNA binding from an intramolecular masking. Unlike the phosphorylated AncEutherian CEBPB, the model of the AncTherian CEBPB phosphorylated at ancestral sites three and four assumes the repressed, transcriptionally inactive, conformation (Fig. 3a). These structural models suggest the differential response of AncEutherian and AncTherian proteins to GSK-3β stimulation results from the acquisition of different conformational states upon phosphorylation.

To test if GSK-3β-mediated phosphorylation induces a conformation change in CEBPB, we immunoprecipitated AncEutherian and AncTherian proteins from BIO-treated or control HeLa cells and incubated the purified protein with 20S proteasome, which degrades proteins with extended unstructured regions[30]. We found that the AncEutherian CEBPB isolated from BIO-treated cells was less efficiently degraded than protein from untreated cells, suggesting that inhibiting GSK-3β shifts the conformation of the AncEutherian CEBPB to the inactive state, which prevents digestion by the 20S proteasome (Fig. 3b). In contrast, AncTherian CEBPB isolated from BIO-treated cells was more efficiently degraded than protein isolated from untreated cells, suggesting that inhibiting GSK-3β shifted the conformation of the AncTherian CEBPB to the active state (Fig. 3b).

To test if GSK-3β-mediated phosphorylation of site five is important for inducing a conformation change in AncEutherian CEBPB, we immunopurified transiently expressed wild-type and site five back-mutated AncEutherian CEBPB in HeLa cells, and either *in vitro* phosphorylated the protein with MAPK/GSK-3β kinases or dephosphorylated it with calf intestinal alkaline phosphatase. We found that *in vitro* phosphorylated wild-type AncEutherian CEBPB was nearly completely degraded by the 20S proteasome, but that back mutation at site five and dephosphorylation protected the protein from degradation (Fig. 3c). These results suggest that in the ancestral state (likely) phosphorylation of sites three and/or four flips the conformation of CEBPB to the closed, transcriptionally repressed state. In contrast, in the derived-state phosphorylation at site five flips the switch to the open, transcriptionally active conformation (Fig. 3d).

These results uncover a potentially important mechanism of gene regulatory evolution and suggest that evolution can affect transcription factor functions in a cell-type-dependent way, probably avoiding deleterious pleiotropic effects on functions in different cellular and developmental contexts. Thus, modification of phosphorylation sites

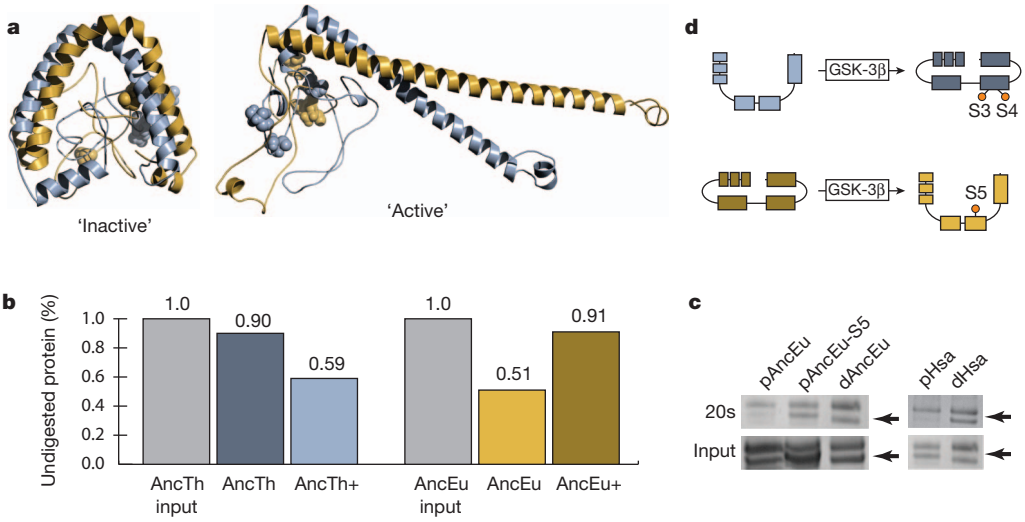

Figure 3 | Phosphorylation induces a conformation change in CEBPB. **a**, Structural models of unphosphorylated and phosphorylated ancestral therian (blue) and eutherian (yellow) CEBPB proteins. The model of the phosphorylated ancestral eutherian CEBPB is in the 'active' conformation whereas the model of the unphosphorylated ancestral therian CEBPB is in the 'active' conformation. Sites three, four and five are shown as spheres. **b**, Ancestral therian (AncTh) and ancestral eutherian (AncEu) CEBPB protein grown in cells treated with the GSK-3β inhibitor BIO (+) or control (−) media are differentially digested by the 20S proteasome. The percentage of undigested protein after a 30-min digestion with the 20S proteasome is shown relative to input. **c**, Phosphorylation of site five induces a conformation change in the ancestral eutherian CEBPB. Incubation of *in vitro* phosphorylated human (pHsa) and ancestral eutherian (pAncEu) CEBPB with purified 20S proteasome leads to protein degradation whereas dephosphorylated proteins (dHsa, dAncEu) are protected from degradation. Back mutation of site 5 (pAncEu-S5) decreases susceptibility of the *in vitro* phosphorylated protein to degradation. Arrows indicate CEBPB bands in Coomassie-blue-stained gel. **d**, Models of the conformational phospho-switch in CEBPB. Phosphorylation of the ancestral therian CEBPB (top) at sites three (S3) and four (S4) by GSK-3β flips the switch from the 'active' conformation (light blue) to the 'inactive' conformation (dark blue). In contrast, phosphorylation of the ancestral eutherian CEBPB (bottom) at site five (S5) by GSK-3β flips the switch from the 'inactive' conformation (dark yellow) to the 'active' conformation (yellow).

in transcription factors can facilitate gene regulatory innovation by altering kinase-signalling-dependent transcriptional networks.

METHODS SUMMARY

We used maximum-likelihood models implemented in CODEML in the PAML4 package of programs to reconstruct ancestral sequences under a JTT+Γ_4 model and the phylogeny shown in Supplementary Fig. 1. HeLa and human ESCs were transiently transfected with 2 ng of *Renilla* control (pGL4.71), 50 ng of dPRL luciferase reporter (dPRL-332/Luc3) and 200 ng of expression vector. Protein samples for co-immunoprecipitation, 20S assay and proteomics were isolated from the nuclear fraction of cell protein, treated with 50 U DNase (Roche) and 2.5 μg RNase (Roche) for 60 min at room temperature and washed three times with 1 ml of lysis buffer. For *in vitro* phosphorylation assays, immunoprecipitated proteins were either incubated with MAPK/GSK-3β and ATP or calf intestinal phosphatase. For 20S assays, *in vitro* phosphorylated and dephosphorylated protein was incubated for 1 h at 37 °C with 1 μg purified human 20S proteosome (Enzo Life Sciences). Immunoprecipitated proteins were either excised from SDS–polyacrylamide gel electrophoresis gels and processed for nano-LC-ESI-MS/MS or transferred to a PVDF membrane for western blotting.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 October 2010; accepted 28 September 2011.
Published online 13 November 2011.

1. Stern, D. L. Evolutionary developmental biology and the problem of variation. *Evolution* **54,** 1079–1091 (2000).
2. Carroll, S. B. Evolution at two levels: on genes and form. *PLoS Biol.* **3,** e245 (2005).
3. Hsia, C. C. & McGinnis, W. Evolution of transcription factor function. *Curr. Opin. Genet. Dev.* **13,** 199–206 (2003).
4. Lynch, V. J. & Wagner, G. P. Resurrecting the role of transcription factor change in developmental evolution. *Evolution* **62,** 2131–2154 (2008).
5. Wagner, G. P. & Lynch, V. J. Molecular evolution of evolutionary novelties: the vagina and uterus of therian mammals. *J. Exp. Zool. B* **304,** 580–592 (2005).
6. Wagner, G. P. & Lynch, V. J. Evolutionary novelties. *Curr. Biol.* **20,** R48–R52 (2010).
7. Lynch, V. J. *et al.* Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. *Proc. Natl Acad. Sci. USA* **105,** 14928–14933 (2008).
8. Galant, R. & Carroll, S. B. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415,** 910–913 (2002).
9. Ronshaugen, M., McGinnis, N. & McGinnis, W. Hox protein mutation and macroevolution of the insect body plan. *Nature* **415,** 914–917 (2002).
10. Löhr, U., Yussa, M. & Pick, L. *Drosophila* fushi tarazu: a gene on the border of homeotic function. *Curr. Biol.* **11,** 1403–1412 (2001).
11. Chen, L. *et al.* Mouse and zebrafish Hoxa3 orthologues have nonequivalent *in vivo* protein function. *Proc. Natl Acad. Sci. USA* **107,** 10555–10560 (2010).
12. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134,** 25–36 (2008).
13. Prud'homme, B., Gompel, N. & Carroll, S. B. Emerging principles of regulatory evolution. *Proc. Natl Acad. Sci. USA* **104,** 8605–8612 (2007).
14. Wray, G. A. The evolutionary significance of *cis*-regulatory mutations. *Nature Rev. Genet.* **8,** 206–216 (2007).
15. Stern, D. L. & Orgogozo, V. The loci of evolution: how predictable is genetic evolution? *Evolution* **62,** 2155–2177 (2008).
16. Shen, F. *et al.* IL-17 receptor signaling inhibits C/EBPβ by sequential phosphorylation of the regulatory 2 domain. *Sci. Signal.* **2,** ra8 (2009).
17. Spooner, C. J., Guo, X., Johnson, P. F. & Schwartz, R. C. Differential roles of C/EBPβ regulatory domains in specifying MCP-1 and IL-6 transcription. *Mol. Immunol.* **44,** 1384–1392 (2007).
18. Friedman, J. R. *et al.* Orthogonal analysis of C/EBPβ targets *in vivo* during liver proliferation. *Proc. Natl Acad. Sci. USA* **101,** 12986–12991 (2004).
19. Christian, M., Pohnke, Y., Kempf, R., Gellersen, B. & Brosens, J. J. Functional association of PR and CCAAT/enhancer-binding proteinβ isoforms: promoter-dependent cooperation between PR-B and liver-enriched inhibitory protein, or liver-enriched activatory protein and PR-A in human endometrial stromal cells. *Mol. Endocrinol.* **16,** 141–154 (2002).
20. Christian, M. *et al.* Cyclic AMP-induced forkhead transcription factor, FKHR, cooperates with CCAAT/enhancer-binding protein β in differentiating human endometrial stromal cells. *J. Biol. Chem.* **277,** 20825–20832 (2002).
21. Pohnke, Y., Kempf, R. & Gellersen, B. CCAAT/enhancer-binding proteins are mediators in the protein kinase A-dependent activation of the decidual prolactin promoter. *J. Biol. Chem.* **274,** 24808–24818 (1999).
22. Williams, S. C., Baer, M., Dilliner, A. J. & Johnson, P. F. CRP2 (C/EBPβ) contains a bipartite regulatory domain that controls transcriptional activation, DNA binding and cell specificity. *EMBO J.* **14,** 3170–3183 (1995).
23. Kowenz-Leutz, E., Twamley, G., Ansieau, S. & Leutz, A. Novel mechanism of C/EBPβ (NF-M) transcriptional control: activation through derepression. *Genes Dev.* **8,** 2781–2791 (1994).
24. Lee, S., Miller, M., Shuman, J. D. & Johnson, P. F. CCAAT/enhancer-binding protein β DNA binding is auto-inhibited by multiple elements that also mediate association with p300/CREB-binding protein (CBP). *J. Biol. Chem.* **285,** 21399–21410 (2010).
25. Kowenz-Leutz, E., Pless, O., Dittmar, G., Knoblich, M. & Leutz, A. Crosstalk between C/EBPβ phosphorylation, arginine methylation, and SWI/SNF/mediator implies an indexing transcription factor code. *EMBO J.* **29,** 1105–1115 (2010).
26. Metz, R. & Ziff, E. cAMP stimulates the C/EBP-related transcription factor rNFIL-6 to trans-locate to the nucleus and induce c-fos transcription. *Genes Dev.* **5,** 1754–1766 (1991).
27. Zhao, X., Zhuang, S., Chen, Y., Boss, G. R. & Pilz, R. B. Cyclic GMP-dependent protein kinase regulates CCAAT enhancer-binding protein β functions through inhibition of glycogen synthase kinase-3. *J. Biol. Chem.* **280,** 32683–32692 (2005).
28. Trautwein, C., Walker, D. L., Plümpe, J. & Manns, M. P. Transactivation of LAP/NF-IL6 is mediated by an acidic domain in the N-terminal part of the protein. *J. Biol. Chem.* **270,** 15130–15136 (1995).
29. Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32,** W526–W531 (2004).
30. Tsvetkov, P. *et al.* Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins Struct. Funct. Bioinf.* **70,** 1357–1366 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by a grant from the John Templeton Foundation, number 12793 Genetics and the Origin of Organismal Complexity; results presented here do not necessarily reflect the views of the John Templeton Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author Contributions V.J.L. and G.P.W. designed experiments and wrote the manuscript. V.J.L. performed experiments and analysed data. G.M. cloned and sequenced CEBPB genes from non-mammalian species.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to V.J.L. (vincent.j.lynch@yale.edu).

METHODS

Identification and cloning of CEBPB genes. CEBPB genes were identified from BLAST/BLAT searches of whole-genome databases at the National Center for Biotechnology Information, University of California, Santa Cruz, and Ensembl including anole (*Anolis carolinensis*), chicken (*Gallus gallus*), zebra finch (*Taeniopygia guttata*), opossum (*Monodelphis domestica*), wallaby (*Macropus eugenii*), armadillo (*Dasypus novemcinctus*), elephant (*Loxodonta africana*), two-toed sloth (*Choloepus hoffmanni*), cow (*Bos taurus*), dolphin (*Tursiops truncates*), Seba's short-tailed bat (*Carollia perspicillata*), flying fox (*Pteropus vampyrus*), horse (*Equus caballus*), dog (*Canis familiaris*), rat (*Rattus norvegicus*), mouse (*Mus musculus*), rabbit (*Oryctolagus cuniculus*), guinea pig (*Cavia porcellus*), pika (*Ochotona princeps*), marmoset (*Callithrix jacchus*), macaque (*Macaca mulatta*), orang-utan (*Pongo pygmaeus*), gorilla (*Gorilla gorilla*), chimpanzee (*Pan troglodytes*) and human (*Homo sapiens*).

We amplified *CEBPB* by PCR from the genomic DNA of two snakes (*Lichanura trivirgata* and *Eryx jayakari*) and two lizards (*Chalcides chalcides* and *Sceloporus undulatus*) using amniote-specific primers (forward: 5′-CCTTTAAATCCA TGGAAGTGGC-3′; reverse: 5′-GACAAGCACAGCGACGAGTA-3′). PCR products were cloned into the pGEM-T vector (Promega). Cloned PCR products were sequenced in both directions by dideoxy chain termination using BigDye chemistry and an automated sequencer. Multiple colonies (at least four) were sequenced for each species.

Ancestral sequence reconstruction. To estimate the substitution rate of the 49 amino acids within and outside placental mammals, we used CODEML in PAML4.0, the tree shown in Supplementary Fig. 1, and f3×4 codon frequencies. Ancestral sequences were inferred with PAML, which uses maximum likelihood and an empirical Bayes approach to estimate ancestral character states, and the phylogeny shown in Supplementary Fig. 1. The Bayesian posterior probability at each site of the reconstructed therian CEBPB ancestral sequence (AncTherian) was 0.96 and the probability of the eutherian CEBPB (AncEutherian) ancestral sequence was 1.0. Ancestral FOXO1A sequences were also inferred with CODEML using the same species for the CEBPB reconstruction. The Bayesian posterior probability of the reconstructed therian FOXO1A ancestral sequence was 0.96 and the probability of the eutherian FOXO1A ancestral sequence was 0.98.

CEBPB expression vector construction. We amplified CEBPB by PCR from the cDNA of human (*Homo sapiens*), opossum (*Monodelphis domesticus*) and chicken (*Gallus gallus*) using primers to designed specific to each species for LAP, the isoform of CEBPB expressed in ESCs. PCR products were cloned into the mammalian expression vector pcDNA3.1(+)-V5/His (Invitrogen) and verified by sequencing. Deletions were made by PCR and cloned into pcDNA3.1(+)-V5/His. Site-directed mutagenesis was performed with the QuickChange Lightning Site Directed Mutagensis kit (Stratagene).

The ancestral therian and ancestral eutherian genes were synthesized by GeneScript Corporation with human-optimized codon usage and ligated into pcDNA3.1(+)-V5/His. Proper expression and nuclear localization of all tagged CEBPB and FOXO1A genes was checked by western blotting using an anti-V5 antibody and nuclear lysate.

Cell culture and luciferase reporter assays. HeLa cells were grown in DMEM supplemented with 10% FBS. Human ESCs were grown in DMEM supplemented with charcoal stripped 10% FBS and differentiated with the progesterone analogue MPA and cAMP. Cells were transiently transfected with *trans*IT-LT1 (Mirus) according to the manufacturer's protocol with 2 ng of the *Renilla* control vector (pGL4.71) and 50 ng of the dPRL luciferase reporter. Depending on treatment, cells were co-transfected with either 200 ng empty pcDNA3.1(+)-V5/His, 100 ng of the FOXO1A expression vector and 100 ng of one of the CEBPB expression vector. Luciferase expression was assayed 48 h after transfection using the Dual Luciferase Reporter System (Promega). Each experiment was repeated four times, with eight replicates per experiment. GSK-3B was inhibited with 10 nM InSolution GSK-3 Inhibitor IX BIO (EMD Chemicals); control media included 10 nM GSK-3 Inhibitor IX, control, MeBIO (EMD Chemicals). We note that caution should be taken when applying luciferase assay results to the situation *in vivo*; we are currently developing a knockdown/xeno-gene rescue system to test the applicability our results to *in vivo* systems.

Co-immunoprecipitations and proteomics. HeLa and hESC cells were cultured as described above. Before transfection, cells were plated at a density of approximately 4.4×10^6 cells per millilitre on 10-cm plates. The next day, CEBPB-V5/His construct was transfected into HeLa cells with either empty pcDNA3.1(+)-V5/His vector, or Flag–FOXO1A pcDNA3.1 vector using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. A total of 12 mg of DNA was mixed with Lipofectamine (1:3) in 3 ml OptiMEM reduced serum media (Gibco) and incubated for 4 h at 37 °C, before addition of 7 ml DMEM. After 16 h the transfection media were removed and replaced with fresh DMEM, and the cells were incubated an additional 24 h before collection.

After removing DMEM and washing cells twice with PBS, 1 ml ice-cold lysis buffer (20 mM Tris, pH 8.0, 40 mM KCl, 10 mM $MgCl_2$, 10% glycerol, 1% Triton X-100, 1× Complete EDTA-free protease inhibitor cocktail (Roche), 1× PhosSTOP (Roche)) was added to each plate and cells were collected by scraping with a rubber spatula; they were then incubated on ice for 30 min after the addition of 5 M NaCl to a final concentration of 420 mM. Whole-cell lysate was cleared by centrifugation at 10,000*g* for 30 min at 4 °C, and supernatant was transferred to a clean microfuge tube. After equilibrating protein concentrations, 1 ml of sample was mixed with 40 ml of M2 anti-Flag agarose beads (Sigma) pre-washed with TNT buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Triton X-100), and rotated overnight at 4 °C.

The next day, samples were treated with 50 U DNase (Roche) and 2.5 μg RNase (Roche) for 60 min at room temperature, as indicated. Samples were washed three times with 1 ml of Hepes Wash (10 mM Hepes, 150 mM NaCl, 0.5% Triton X-100). Between each wash, agarose beads were collected by centrifugation at 1,000*g* for 1 min at 4 °C and supernatant was removed. After the final wash, beads were re-suspended in elution buffer (500 mM Tris pH 7.5, 1 M NaCl) and incubated for 1 h at 4 °C. For *in vitro* phosphorylation assays, immunoprecipitated proteins were either incubated with MAPK/GSK-3β and ATP (New England Biolabs) or calf intestinal phosphatase (New England Biolabs) according to the manufacturer's protocol. For 20S assays, *in vitro* phosphorylated and dephosphorylated protein was incubated for 30 min at 37 °C with 1 μg purified human 20S proteosome (Enzo Life Sciences) in 40 μl digestion buffer.

Immunoprecipitated proteins were either excised from SDS–polyacrylamide gel electrophoresis gels and sent to the Keck Proteomics centre for further processing and nano-liquid chromatography–electrospray ionization/tandem mass spectrometry-based identification of post-translational modifications, or transferred to a PVDF membrane for western blotting. After transfer to PVDF, the membrane was blocked with TBS Tween20 + BSA (25 mM Tris, 150 mM NaCl, 3% BSA), then incubated with anti-V5-HRP antibody (1:5,000; Invitrogen) for 2 h. Bands were visualized using the ECL SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). Membranes were stripped with Restore Western Blot stripping buffer (Thermo Scientific) and re-probed with anti-Flag-HRP antibody (1:100,000; Sigma).