npg

Tissue-specific targeting of cell fate regulatory genes by E2f factors

LM Julian[1,4], Y Liu[2], CA Pakenham[1], D Dugal-Tessier[1], V Ruzhynsky[1], S Bae[3], S-Y Tsai[3], G Leone[3], RS Slack*,[1] and A Blais*,[2]

Cell cycle proteins are important regulators of diverse cell fate decisions, and in this capacity have pivotal roles in neurogenesis and brain development. The mechanisms by which cell cycle regulation is integrated with cell fate control in the brain and other tissues are poorly understood, and an outstanding question is whether the cell cycle machinery regulates fate decisions directly or instead as a secondary consequence of proliferative control. Identification of the genes targeted by E2 promoter binding factor (E2f) transcription factors, effectors of the pRb/E2f cell cycle pathway, will provide essential insights into these mechanisms. We identified the promoter regions bound by three neurogenic E2f factors in neural precursor cells in a genome-wide manner. Through bioinformatic analyses and integration of published genomic data sets we uncovered hundreds of transcriptionally active E2f-bound promoters corresponding to genes that control cell fate processes, including key transcriptional regulators and members of the Notch, fibroblast growth factor, Wnt and Tgf-β signaling pathways. We also demonstrate a striking enrichment of the CCCTC binding factor transcription factor (Ctcf) at E2f3-bound nervous system-related genes, suggesting a potential regulatory co-factor for E2f3 in controlling differentiation. Finally, we provide the first demonstration of extensive tissue specificity among E2f target genes in mammalian cells, whereby E2f3 promoter binding is well conserved between neural and muscle precursors at genes associated with cell cycle processes, but is tissue-specific at differentiation-associated genes. Our findings implicate the cell cycle pathway as a widespread regulator of cell fate genes, and suggest that E2f3 proteins control cell type-specific differentiation programs by regulating unique sets of target genes. This work significantly enhances our understanding of how the cell cycle machinery impacts cell fate and differentiation, and will importantly drive further discovery regarding the mechanisms of cell fate control and transcriptional regulation in the brain, as well as in other tissues.

The classical cell cycle regulatory pRb/E2f pathway has emerged as an important effector of fate decisions in a number of cell types, including in the brain. Cell cycle dynamics strongly influence neural precursor cell (NPC) maintenance and neurogenesis,[1–4] and gain- or loss-of-function studies have demonstrated key roles for cell cycle proteins, including the E2f family, in NPC fate decisions.[3–16] E2f3 is required for proper cortical migration of neurons and to maintain the balance between NPC self-renewal, proliferation and differentiation, and its loss disrupts long-term neurogenesis and cortical function; E2f1 deficiency impairs NPC proliferation, and E2f4 deficiency leads to inhibition of NPC self-renewal and severe defects in telencephalic development.[6,8–10,17]

A pivotal question is whether cell fate control by the pRb/E2f pathway is largely a consequence of cell cycle regulation, or due to direct regulation of cell fate-associated genes. We recently found that loss of E2f3a and E2f3b leads to opposing defects in NPC maintenance and differentiation.[8] The fact that

this occurred without affecting cell cycle dynamics strongly suggests that fate control by E2fs is not secondary to cell cycle regulation. In addition, a number of key cell fate genes and pathways have been identified as E2f-regulated targets driving E2f-dependent fate decisions in NPCs. These include the neurogenesis and migration genes Dlx1/Dxl2 and Neo1 (Neogenin),[18,19] the growth factor fibroblast growth factor 2 (Fgf2),[17] the pluripotency and self-renewal factor Sox2,[8] and the Notch/Hes[11] and Sonic Hedgehog pathways.[10] Together, these findings demonstrate a direct role for pRb/E2f at cell fate-associated genes, but the extent of this interaction is unknown. Many of these studies focused on single pRb or E2f factor knock-out models, and because E2fs exhibit extensive redundancy in their biological functions and genomic binding patterns,[20,21] it is likely that more E2f target genes relevant to cell fate exist than those that have been reported. Understanding the full regulatory potential of the cell cycle machinery in the brain therefore necessitates an appreciation

[1]Neuroscience program, Department of Cellular and Molecular Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada; [2]Department of Biochemsitry, Microbiology and Immunology, Ottawa Institute of Systems Biology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada and [3]Solid Tumor Biology Program, Department of Molecular Virology, Immunology, and Medical Genetics and Department of Molecular Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA
*Corresponding author: A Blais, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada. Tel: +613 562 5800 extn. 8253; Fax: +613 562 5452; E-mail: Alexandre.Blais@uottawa.ca
or RS Slack, Department of Cellular and Molecular Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada. Tel: +613 562 5800 extn. 8458; Fax: +613 562 5403; E-mail: rslack@uottawa.ca
[4]Current address: Sprott Centre for Stem Cell Research, Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada.
Abbreviations: ChIP, chromatin immunoprecipitation; Ctcf, CCCTC binding factor; E (ie. E14.5), embryonic day; E2f, E2 promoter binding factor; Fgf, fibroblast growth factor; GO, gene ontology; MB, myoblast; NPC, neural precursor cell; TSS, transcriptional start site; WT, wild-type
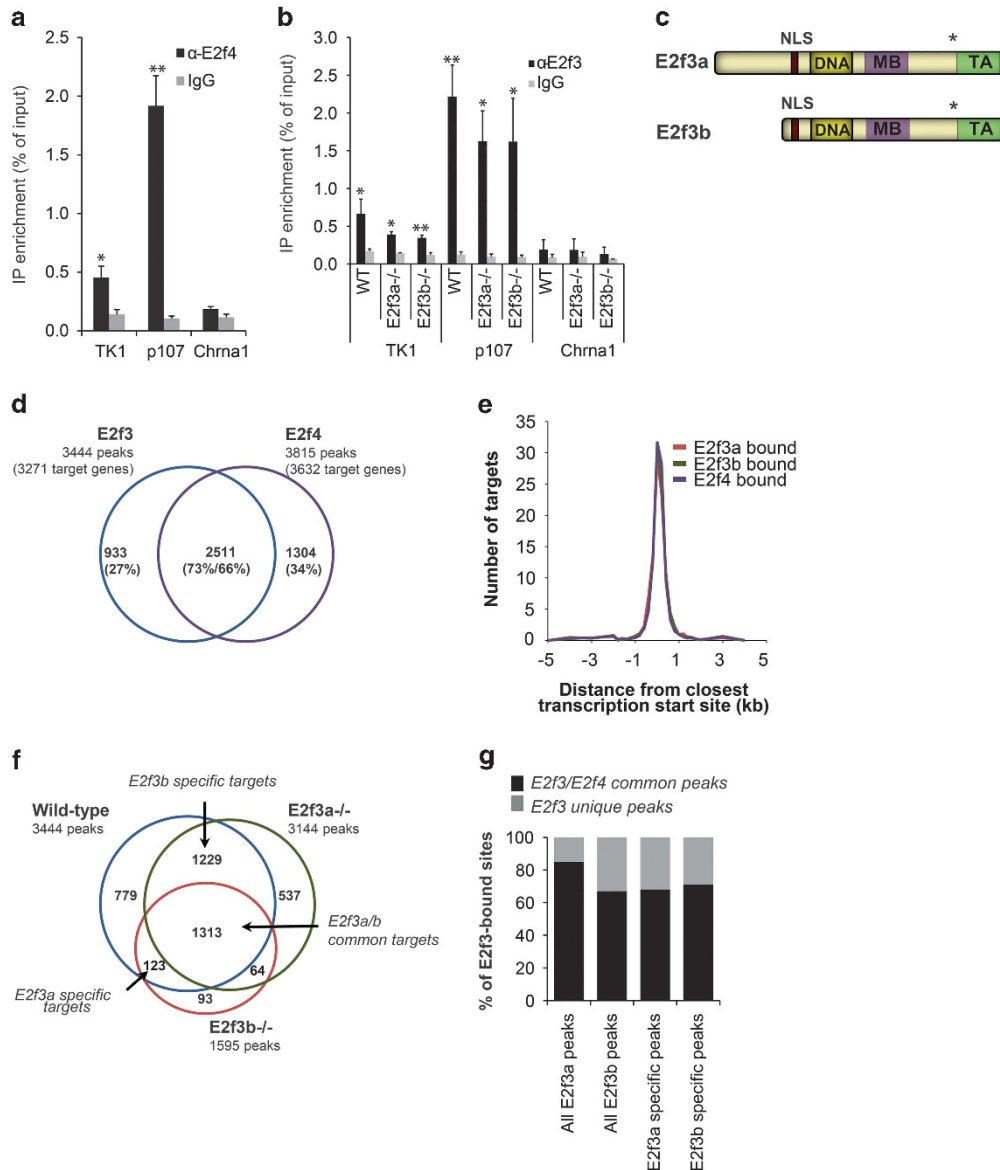
Figure 1 Specificity and redundancy among E2f3 and E2f4 genomic binding sites in neural precursor cells. (**a**) Quantitative ChIP (qChIP) analysis of E2f4 binding at the promoters of previously identified E2f targets (thymidine kinase (TK1) and p107) and a negative control gene (Chrna1) in WT neurospheres. Normal rabbit IgG is used as a non-specific antibody control. E2f4 is significantly enriched over rabbit IgG at positive control genes only. *$P < 0.05$, **$P < 0.01$. Graphed are the averages of four to six biological replicates, $\pm$ S.E.M. (**b**) qChIP analysis to assess E2f3 binding at TK1, p107 and Chrna1 promoters only, with similar enrichment in WT and knock-out cells. (*$P < 0.05$, **$P < 0.01$). Graphed are the averages of three to four biological replicates, $\pm$ S.E.M. (**c**) Schematic diagram of E2f3a and E2f3b protein structure. Both isoforms share common structural domains: DNA, DNA-binding domain; MB, marked box domain; NLS, nuclear localization sequence; TA, transcriptional activation domain, The region recognized by the α-E2f3 antibody is denoted by '*'. (**d**) Venn diagram demonstrating the degree of overlap between E2f3 and E2f4 peaks discovered by ChIP–chip. The number of peaks in each category is shown, with the percent of all E2f3 or E2f4 peaks in that category in parentheses. (**e**) Distribution of E2f3a-, E2f3b- and E2f4-bound sites relative to the nearest TSS. The three binding data sets were categorized into 200 bp bins, and the percent of bound sites in each category (indicated in upper right corner) within each bin range is plotted. (**f**) Venn diagram demonstrating the overlap between E2f3 peaks identified by ChIP–chip in WT, E2f3a $-/-$ and E2f3b $-/-$ neurospheres. Peaks identified as specific or common to E2f3a and E2f3b are indicated. (**g**) Analysis of the percent of E2f3 peaks in the indicated categories that overlap with E2f4 peaks in NPCs

of the repertoire of E2f target genes in NPCs. Genome-wide targets have been reported for E2f4 (and E2f1) in a number of immortalized and transformed cell lines,[20,22,23] and for E2f3 in immortal mouse myoblasts, myotubes and embryonic fibroblasts.[24,25] E2f targets have, however, not been identified on a genome-wide level in any neural cells, nor in any primary cell types.

Here we used an unbiased approach to show that E2f3a, E2f3b and E2f4 target a large and diverse network of cell fate regulatory genes, with substantial overlap in their binding sites. We further show that E2f3 isoforms demonstrate extensive tissue specificity in the genes that they target, positioning them as potential regulators of specialized, cell type-specific differentiation programs. Our results highlight the

cell cycle machinery as a pervasive regulator of cell fate decisions, and provide a comprehensive tool to fuel further discovery.

Results

E2f3&4 promoter occupancy in neural precursor cells.
We determined the genomic binding sites of E2f3 and E2f4 in NPCs derived from telencephalic tissue at embryonic day 14.5 (E14.5), a peak stage of cortical neurogenesis. As E2f binding sites are typically found within close proximity of a transcriptional start site (TSS)[8,17,20,24,26] we focused our analysis on promoter regions by coupling chromatin immunoprecipitation (ChIP) using antibodies towards E2f3 and E2f4 (Figure 1a and b) with proximal promoter DNA microarrays (details in Supplementary Text S1). We have validated the specificity of these antibodies previously[8] and here (Supplementary Figure S1 and S2). We first identified E2f3-binding sites in wild-type (WT) cells using an antibody that binds a C-terminal epitope common to both E2f3a&b (Figure 1c). Analysis of these sites based on the gene whose TSS is closest to each peak revealed 3444 peaks enriched for E2f3, corresponding to 3271 unique target genes (Figure 1d, Supplementary Table S1). Importantly, E2f3 peak sequences were enriched compared with control promoter regions for the E2f consensus motif, with the highest ranked peaks exhibiting the greatest motif enrichment (discussed below), underscoring the validity of our approach. The data were similar for E2f4, with 3815 peaks and 3632 target genes (Figure 1d, Supplementary Table S1). In total, we identified 5459 unique E2f-bound peaks in NPCs, and found that 28% of all proximal promoter regions are bound by at least one E2f3/4 factor. Furthermore, while we surveyed genomic regions encompassing − 5 to +3 kb relative to all known TSSs (Supplementary Figure S3), 88–90% of enriched peaks fell within 600 bp upstream and 600 bp downstream of a TSS (Figure 1e). Interestingly, while a large fraction of E2f3 and E2f4 binding sites overlap, approximately one-third of the binding sites for each protein are unique (Figure 1d, Supplementary Table S2), suggesting the existence of factor-specific target genes.

Based on their binding profiles in muscle cells,[24] we hypothesized that E2f3a&b may bind unique target genes in NPCs. Thus, we performed ChIP–chip for E2f3 on chromatin from E2f3a − / − and E2f3b − / − NPCs. This approach allowed us to identify genomic regions bound by E2f3a alone (present in both WT and E2f3b − / − , absent in E2f3a − / − cells), E2f3b alone (present in WT and E2f3a − / − , absent in E2f3b − / − cells) and by both E2f3a and E2f3b (present in all three conditions) with the use of a single antibody (Figure 1f). Importantly, we have previously demonstrated that loss of either E2f3 isoform in NPCs does not result in altered expression or a perceivable change in DNA-binding capacity of the other isoform, nor of other E2f factors.[8] We identified 3144 enriched peaks (2979 genes) in E2f3a-deficient cells, indicating that E2f3b is present at the majority of E2f3-bound sites in NPCs. In contrast, we found only 1595 peaks (1518 genes) in E2f3b − / − cells, suggesting that E2f3a is present at a smaller subset of E2f3-bound sites (Figure 1f, Supplementary Table S3). The 779 peaks enriched for E2f3
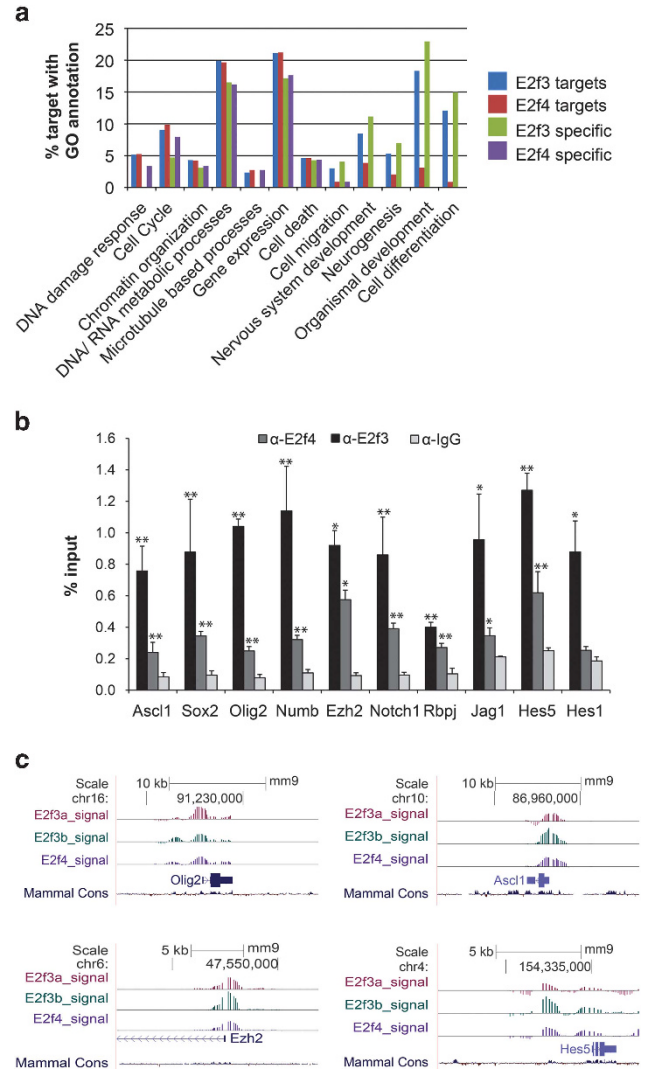
Figure 2 E2f3 and E2f4 bind an overlapping set of gene promoters associated with fundamental NPC fate decisions. (**a**) Gene Ontology (GO) analysis of E2f3 and E2f4 target genes, expressed as the percentage of all target genes in each group (groups indicated to the right of the graph) with a particular GO annotation. Target genes were identified genes whose TSS is closest to each E2f peak. (**b**) qChIP analysis validates E2f3 and E2f4 enrichment at binding sites identified by ChIP–chip corresponding to a panel of genes involved in neural precursor cell fate decisions. IgG is used as a negative control. Both E2f3 and E2f4 are enriched at all analyzed sites, except for that corresponding to the gene Hes1, which is uniquely detected by anti-E2F3. *$P < 0.05$, **$P < 0.01$. The average of at least three biological replicates is shown, ± S.E.M. (**c**) Genomic binding profiles for E2f3a (E2f3 ChIP–chip in E2f3b − / − neurospheres), E2f3b (E2f3 ChIP–chip in E2f3a − / − neurospheres) and E2f4 (ChIP–chip in wild-type cells), at selected loci. At each locus shown, all three E2f family members were significantly enriched (E2f identity indicated to the left). The level of phylogenetic conservation across 17 species of mammals is shown along the bottom of each image ('Mammal Cons'). The majority of E2f3 peaks are located within highly conserved regions, and most promoters are bound by multiple E2fs with highly overlapping peaks. Peaks are generally found overlapping or directly proximal to the TSS. The black arrows specify the location and direction of transcription initiation

in WT cells only (Figure 1f, Supplementary Table S3) probably represent a class of loci that require a maximum amount of E2f3 protein for optimal ChIP detection. This claim is supported by the fact that 78% of these E2f3 WT 'unique'

Table 1 Quantification of select biological processes enriched among genes targeted by E2f3 and E2f4

GO term	Gene count (E2f3; E2f4)	% Of all genes (E2f3; E2f4)	P-value (E2f3)	Fold enrichment (E2f3)
DNA damage response	178; 201	5.2; 5.3	8.68E − 55	3.53
Cell cycle	312; 376	9.1; 9.9	6.51E − 79	3.05
Chromatin organization	148; 162	4.3; 4.3	6.43E − 29	2.73
DNA/RNA metabolic processes	685; 809	19. 9; 19.7	9.10E − 118	2.38
Microtubule-based processes	81; 105	2.4; 2.8	1.82E − 10	2.37
Gene expression	724; 809	21.1; 21.2	7.13E − 123	2.35
Cell death	159; 177	4.6; 4.6	2.89E − 13	1.93
Cell migration	104; 11	3.0; 0.9	6.63E − 07	1.87
Nervous system development	291; 148	8.5. 3.9	1.37E − 17	1.71
Neurogenesis	184; 79	5.3; 2.1	5.78E − 08	1.62
Organismal development	631; 119	18.3; 3.1	8.97E − 31	1.56
Cell differentiation	405; 34	12.1; 0.9	6.16E − 17	1.54

E2f, E2 promoter binding factor
Given are the number and percent of genes bound by either E2f3 or E2f4 belonging to particular GO annotations, as determined by GREAT.[57] Fold enrichment is calculated based on the number of E2f3-bound genes with a given annotation compared with the number of genes in the background set with that annotation. The file used to generate genes for 'background' comparisons in these calculations was all known mouse promoter regions that were surveyed on our DNA microarrays

sites overlap with peaks in E2f3a − / − and E2f3b − / − cells for which an E2f3-binding signal was obtained but did not meet our FDR requirement (Supplementary Figure S4). At E2f-bound promoters, while E2f3a is enriched at only a subset of E2f3b-bound sites, E2f4 peaks overlap significantly with both E2f3a and E2f3b peaks (Supplementary Figure 1G). Thus, at least two E2f family members can bind the majority of our E2f-bound sites.

E2fs bind a network of genes involved in cell fate decisions. We performed gene ontology (GO) analysis to identify the biological processes with which E2f target genes are associated in NPCs. We found a high enrichment for both E2f3 and E2f4 targets in well-characterized E2f-regulated functions, including the DNA damage response, cell cycle regulation, chromatin organization, gene expression and cell death[23,24,26–31] (Figure 2a, Table 1). Previously established E2f target genes identified in NPCs, most common to both E2f3&4, include *Pcna, Ccne, Top2a, Cdc2a, Cdc25a, Mcm4, Rad51, Brca1/2, Ezh2* and pRb and E2f family members[24,26,32–34] (complete lists in Supplementary Tables S1 and S2). In agreement with pRb/E2f family loss-of-function studies, we also observed a strong enrichment in processes related to differentiation and development, including those specific to the nervous system (Figure 2a, Table 1). More surprisingly, however, identification of the target genes themselves (Figure 3a, for example, Supplementary Tables S1 and S2) revealed that E2fs are bound to the promoters of hundreds of genes that control cell fate decisions (Table 1). Furthermore, these targets are functionally diverse, including genes associated with key growth factor and developmental signaling pathways, chromatin modifiers and transcription factors that together regulate stem cell identity, self-renewal and differentiation.

We observed that E2f3&4 share a number of NPC cell fate genes as targets. Examples include genes or members of pathways that have previously been described as targets of the pRb-E2f family in the brain, such as *Sox2*[8,35] and members of the Notch/Hes (*Notch1, Hes5, Jag1, Rbpj* and *Numb*)[11,14] and Fgf pathways (*Fgfr2* and *Fgfr3*).[17] Additional genes include components of the Wnt (*Wnt6, Fzd5, Dll4, Gsk3b,*

Apc and *Tcf3*) and Tgf-β signaling pathways (*Tgfbr1, Bmp7, Smad4&7* and *Acvr2b*), the transcription factors *Ascl1, Pax6* and *Lhx2*, the chromatin regulators *Dnmt1, Mbd1* and members of the Polycomb/Trithorax families (*Ezh2, Eed, Epc1* and *Mll1*). We confirmed binding of E2f3 and E2f4 to the promoters of a group of target genes that are key cell fate regulators in NPCs, using quantitative ChIP (Figure 2b). We note that the majority of these genes show enrichment for both E2f4 and at least one E2f3 isoform, typically with tightly overlapping peaks, and that these binding sites largely correspond with evolutionarily conserved regions (Figure 2c).

Despite this overlap, quantitative GO analysis surprisingly revealed that the functional categories related to development and differentiation are more highly enriched among E2f3 than E2f4 targets (Figure 2a, Table 1). In fact, E2f3 binds 12-fold more differentiation-related genes (405 versus 34) and twice as many nervous system development and neurogenesis-related genes than does E2f4 (Table 1). In addition, migration-related genes, such as those involved in actin dynamics and axon guidance, are also more heavily targeted by E2f3 (Table 1, Supplementary Table S4). Cell fate-related genes unique to E2f3 primarily include additional transcription factors, chromatin regulators and family members of the core pathways represented among E2f3/E2f4 common targets (such as *Mecp2, Atrx, Wnt5a, Vegf, Hes1, Egfr* and *Fgfr1*).

Genes targeted by E2f3a and E2f3b control common cellular processes in NPCs. GO analysis demonstrated that E2f3a and E2f3b target genes are enriched for similar functional categories in NPCs whether they are common or unique to one isoform (Figure 3b, Supplementary Table S5). E2f3a-specific targets are somewhat less enriched for certain functions, including 'cell cycle' and 'embryo development', but overall our analysis reveals that functions as wide-ranging as the DNA damage response, gene expression and neurogenesis are relatively equally enriched among E2f3a- and E2f3b-specific and common target genes.

To focus on the genes that are likely directly related to NPC fate decisions, we grouped E2f3 target genes belonging to the 'neurogenesis' and 'nervous system development' categories. About 70% of this group of genes is targeted either exclusively
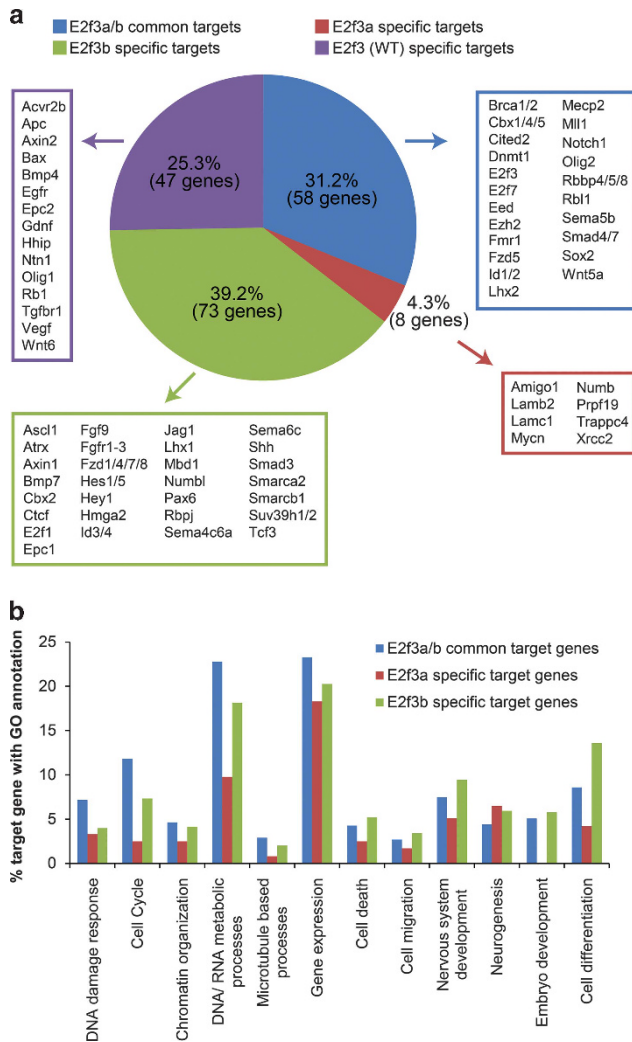
a

- ■ E2f3a/b common targets
- ■ E2f3a specific targets
- ■ E2f3b specific targets
- ■ E2f3 (WT) specific targets

Acvr2b	
Apc	
Axin2	
Bax	
Bmp4	
Egfr	
Epc2	
Gdnf	
Hhip	
Ntn1	
Olig1	
Rb1	
Tgfbr1	
Vegf	
Wnt6	

25.3% (47 genes)
31.2% (58 genes)
39.2% (73 genes)
4.3% (8 genes)

Brca1/2	Mecp2
Cbx1/4/5	Mll1
Cited2	Notch1
Dnmt1	Olig2
E2f3	Rbbp4/5/8
E2f7	Rbl1
Eed	Sema5b
Ezh2	Smad4/7
Fmr1	Sox2
Fzd5	Wnt5a
Id1/2	
Lhx2	

Amigo1	Numb
Lamb2	Prpf19
Lamc1	Trappc4
Mycn	Xrcc2

Ascl1	Fgf9	Jag1	Sema6c
Atrx	Fgfr1-3	Lhx1	Shh
Axin1	Fzd1/4/7/8	Mbd1	Smad3
Bmp7	Hes1/5	Numbl	Smarca2
Cbx2	Hey1	Pax6	Smarcb1
Ctcf	Hmga2	Rbpj	Suv39h1/2
E2f1	Id3/4	Sema4c6a	Tcf3
Epc1			

b

Figure 3 Genes involved in nervous system development and neurogenesis are targeted by both E2f3 isoforms. (**a**) The number of E2f3 target genes with the GO classifications of 'nervous system development' or 'neurogenesis' was determined, and the percentage of these genes that belong to specific groups of E2f3a and E2f3b targets are shown. Examples of genes belonging to each category of E2f3 targets are listed. (**b**) E2f3 isoforms control common cellular functions. GO analysis of E2f3a and E2f3b common and specific target genes, expressed as the percentage of all target genes in each group with a particular GO annotation (GO annotations indicated at the bottom of the graph). Target genes were identified genes whose TSS is closest to each E2f peak

E2f3 and Ctcf are co-enriched at the promoters of neural specific target genes. We analyzed our E2f3-bound regions for the occurrence of binding sites of other transcription factors, to uncover combinatorial regulation partners of E2f3. As expected, we found a strong over-representation of sequence motifs corresponding to E2F factors, as well as other cell cycle-related factors, such as NF-Y and Sp1 (data not shown).[36] Less expectedly, we uncovered a specific enrichment of CCCTC binding factor (Ctcf) binding motifs across E2f3 peaks (Figure 4a). Exploiting data sets of genomic Ctcf sites identified by ChIP-Seq in mouse neuronal tissues,[29] we confirmed a remarkable overlap between binding sites of E2f3 and Ctcf in mouse brain at E14.5. E14.5 mouse brain is a tissue undergoing active neurogenesis and NPC expansion, and is therefore amenable to comparison with our NPCs.[37] We found that 45% of E2f3 target promoters are indeed jointly bound by Ctcf, a significant enrichment (1.46-fold) compared with the overlap expected based on all promoters surveyed by our microarrays (Figure 4b, Supplementary Table S6). Furthermore, the enrichment of either E2f or Ctcf sequence motifs was strongest at peaks where both factors are bound (Table 2). Distinguishing between E2f3- and E2f4-specific targets revealed that the association with Ctcf is strongest for E2f3 (Figure 4b).

Surveying our E2f3-bound promoters with Ctcf peaks from various brain tissues revealed that the E2f3–Ctcf association is strongest in tissues undergoing neurogenesis (E14.5 brain and olfactory bulb), and is substantially lower in the less neurogenic adult cortex and cerebellum (Figure 4c). In sharp contrast to NPCs, the E2f3–Ctcf association is not enriched in myoblasts (MBs) and is in fact slightly depleted at MB-specific E2f3 targets (Figure 4d). These observations suggest that genes bound by both E2f3 and Ctcf are more likely to be involved in neurodevelopmental processes. GO analysis confirmed that while E2f3 targets are equally represented for 'cell cycle', 'chromatin organization' and 'gene expression' whether or not Ctcf is also bound, targets at which both E2f3 and Ctcf are bound are uniquely enriched for functions and signaling pathways that impact cell fate-related processes, many with important roles during neurogenesis (Figure 4e). In addition, genes targeted by E2f3 have a stronger tendency to be targets of Ctcf than those not targeted by E2f3, and this association is most predominant for genes involved in functions related to brain development or neuron generation (Figure 4f). Finally, we observed that E2f3-bound promoters that are also bound by Ctcf are characterized by two additional properties associated with developmentally important genes: they are more highly phylogenetically conserved (Figure 4g) and are predicted to associate with a larger number of distal enhancers (Figure 4h, Supplementray Table S7). Together, these data suggest that a strong overlap between E2f3 and Ctcf may be a feature specific to neural tissue and may at least partially account for E2f3-dependent target gene binding in NPCs.

E2f3 is enriched at transcriptionally active genes in NPCs and exhibits tissue-specific binding. Generally, E2f3a is considered a transcriptional activator, while E2f3b and E2f4 are thought to function predominantly as repressors.[38–40] We

by E2f3b (39.2%) or by both E2f3a and E2f3b (31.2%), while a minority is bound specifically by E2f3a (3.5%) or by E2f3 in WT cells only (25.3%) (Figure 3a). Notably, among the genes identified as specific to E2f3b are members of the Fgf signaling pathway, which we have previously shown is regulated by E2f3.[17] On the other hand, as was the case for E2f3&4 common targets, we found that many genes with established roles as regulators of NPC fate decisions are shared between E2f3a&b (see Figure 3a for examples of E2f3 isoform specific and common genes). These data, combined with those for E2f4 (Figures 1d and g), further demonstrates that the majority of E2f targets in NPCs are bound and likely regulated by multiple E2f proteins.

evaluated the relative mRNA levels of genes in E14.5 brain tissue whose promoters are bound by different combinations of E2f3a, E2f3b and E2f4,[37] and surprisingly found that all categories of E2f-bound genes exhibit higher expression levels than average, compared with all genes surveyed on our microarrays. This observation suggests a transcriptional activation role for all three E2fs in NPCs (Figure 5a), which is

corroborated by the strong association we observed between bound E2f3 and the presence of chromatin marks that correlate with gene activation (Supplementary Figure S5). Of note, we found that E2f3b-specific genes collectively showed significantly lower expression compared with both E2f3a&b-common and E2f3a-specific targets, demonstrating that the presence of E2f3a is associated with higher levels of gene
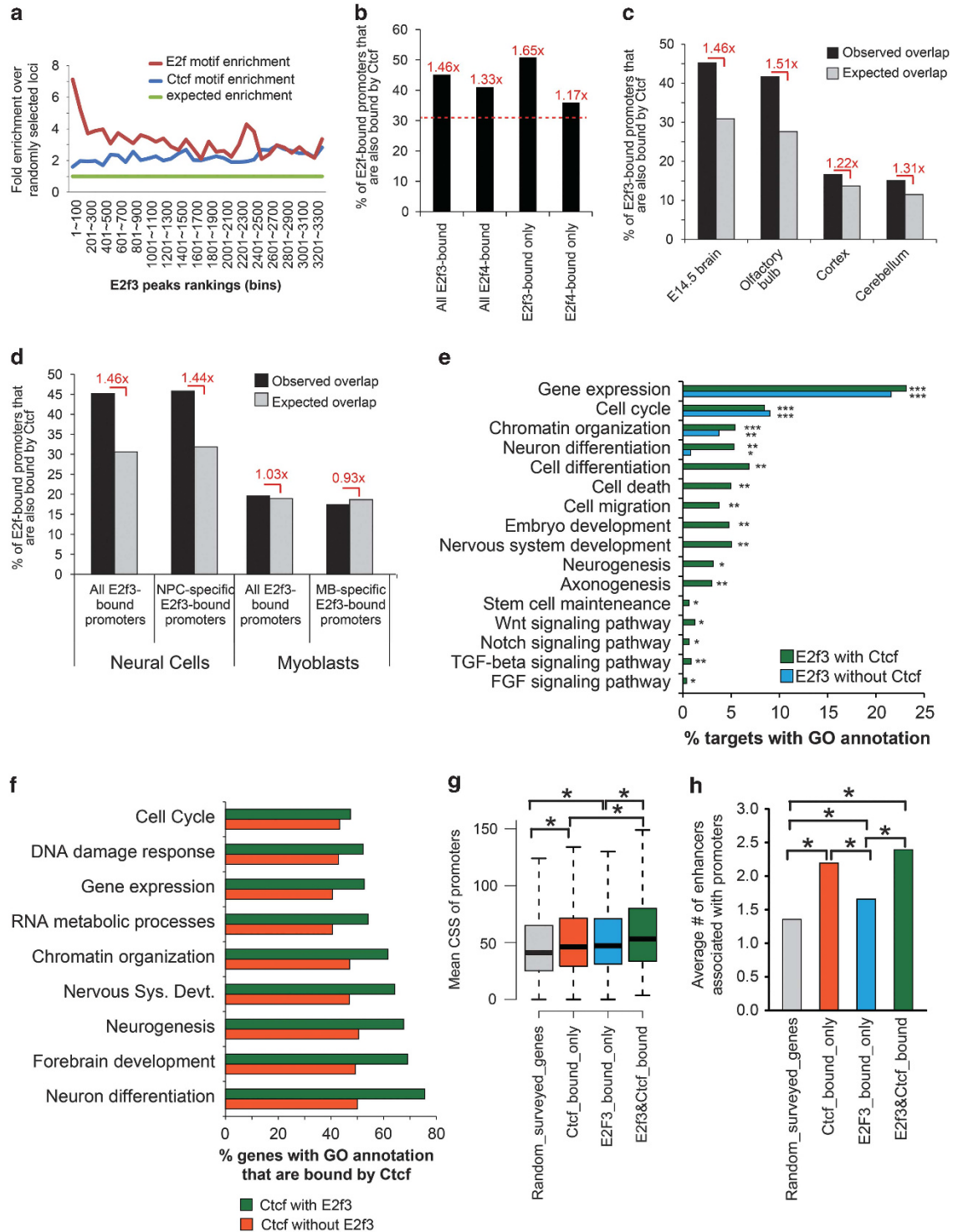
expression compared with E2f3b (Figure 5a). Furthermore, genes whose promoters are bound by E2f3 in NPCs have higher-than-average expression patterns in many of the tissue types examined, but it is in E14.5 brain and other brain tissues that E2f3 binding is most associated with high gene expression (Figure 5b). Quantitative RT-PCR in E2f3b +/+ and E2f3b −/− NPCs validated that E2f3 is important for expression of cell fate genes (Figure 5c). While overall the effects of E2f3b loss on gene expression were modest and some genes were unaffected (likely due to redundant binding by other E2f factors), half of these genes showed altered expression below or close to a *P*-value of 0.05. In addition, the majority of these altered genes showed reduced as opposed to increased expression in the absence of E2f3b, further suggesting that E2f3b functions predominantly (although not exclusively) as a transcriptional activator. In contrast, the NPC target genes of E2f3 are expressed at levels much closer to the average in some tissues, including heart, liver and testes. By analyzing a genome-wide atlas of gene expression in mouse tissues, we identified a cluster of over 400 E2f3 targets that are highly expressed predominantly in neuronal tissues (Supplementary Figure S6, Supplementary Table S9). These observations suggest that E2f3-bound sites are differentially transcribed depending on the cell type, and/or that E2fs exhibit significant tissue specificity in their genomic binding patterns.

Table 2 Enrichment of TF motifs across promoters bound by E2f3, Ctcf or both E2f3 and Ctcf

Factor(s) bound at promoters	E2f motif enrichment	Ctcf motif enrichment
E2f3 only	2.00	1.10
Ctcf only	1.07	1.55
E2f3+Ctcf	2.23	1.81

Abbreviations: Ctcf, CCCTC binding factor; TF, transcription factor
The enrichment is given as a proportion over the frequency of each motif among a randomly selected group of promoters surveyed in our ChIP–chip experiments

To assess tissue-specific binding, we compared our E2f3 peaks from NPCs with those previously identified in MBs.[24] In this analysis, we compared only those peaks that were discovered using the same antibody (sc-878) in each cell type and that fell within genomic regions that were surveyed in both experiments. We found a surprisingly low degree of overlap, with just 30% of MB genes also identifying as targets in NPCs (Figure 6a). Strikingly, GO analysis revealed a strong conservation of classical E2f processes between the two cell types (cell cycle, gene expression, DNA damage response, chromatin organization and cell death). In contrast, tissue-specific E2f3 targets were enriched for developmentally-related functions, and these categories are poorly represented within the group of genes targeted by E2f3 in both NPCs and MBs (Figure 6b, Supplementary Table S10). Specifically, E2f3 targets are uniquely enriched in NPCs for 'cell differentiation', 'nervous system development' and 'neurogenesis', while MB-specific targets are primarily enriched for 'cell adhesion', 'response to wounding' and 'skeletal system development'. These findings demonstrate both a conservation of 'classical' E2F-dependent functions between cell types, which surprisingly occurs through a relatively small proportion of E2f3-binding sites, as well as extensive cell type specificity in terms of both target gene binding and cellular functions. Together, our comparative analyses reveal the surprising finding that E2f3 is positioned to control distinct differentiation and cell fate regulatory pathways in different mammalian cell types.

Discussion

The cell cycle machinery is a pivotal regulator of brain development and function by influencing key cell fate decisions, typically via E2f transcription factor activity.[5,6,8,9,17,18,41–43] Thus through our identification of E2f3&4-bound promoters in NPCs, we have significantly expanded our understanding of how cell cycle regulators can direct cell fate control. While this study provides a number of important insights, we offer two key conceptual findings: E2f transcription factors are poised as widespread regulators of cell fate-associated genes in NPCs, establishing a pervasive direct role for the cell cycle machinery

Figure 4 E2f3 and Ctcf are co-enriched at nervous system-related genes. (**a**) Enrichment of E2f and Ctcf position weight matrices (PWM) within E2f3 peaks discovered in NPCs, as compared with the frequency of their occurrence in randomly selected surveyed loci. The genomic loci enriched for E2f3 were ranked in decreasing order of ChIP–chip enrichment and were binned into groups of 100 target loci. (**b**) The percentage of E2f3- and/ or E2f4-bound regions (E2f-bound category indicated along the *x*-axis) in NPCs at which Ctcf is also bound (in E14.5 brain) was calculated. Numerical values above each data set indicates the fold change of the observed over expected percentage overlap. The expected overlap (30.8%) is denoted by a dashed line, and represents the number of all surveyed promoters at which Ctcf is bound. (**c**) The percentage of promoter regions at which E2f3 is bound in NPCs that also contained Ctcf in different brain tissues (indicated along the *x*-axis) was quantified and is expressed as 'observed overlap'. The 'expected overlap' is the percentage of all known promoter regions surveyed in our ChIP–chip experiments that contain a Ctcf peak. Numerical values above each data set indicates the fold change of the observed over expected percentage overlap. (**d**) The percentage of E2f3-bound promoters at which Ctcf is also bound was calculated separately for neural cells and myoblasts. We show, as the 'expected overlap', the percentage of E2f3-bound promoters containing Ctcf for (1) all E2f3-bound genes in the specified cell type, (2) E2f3-bound genes specific to that cell type versus the other, and 3) E2f3-bound genes common to both cell types. For neural cells, E2f3-binding sites were sourced from our data, and the Ctcf data were from E14.5 brain tissue;[29] for myoblasts E2f3 data were from ref. 24 and the Ctcf data were reported by the Wold lab (ENCODE project). The 'expected overlap' represents the percentage of gene promoters surveyed in each cell type that contain Ctcf, for either all genes surveyed (the first data set for each cell type) or genes surveyed in both cell types (the last two data sets for each group). (**e**) Comparative GO analysis of E2f3-bound target genes that are bound by Ctcf versus those not bound by Ctcf in E14.5 brain. The percentage of E2f3 targets in each group with specific GO annotations (indicated on *y*-axis) was determined by GREAT. *P*-values are indicated (*$P < 1.0E - 4$, **$P < 1.0E - 10$ and ***$P < 1.0E - 20$). (**f**) The percent of surveyed genes whose promoter region is bound by Ctcf in E14.5 brain that belong to various GO categories (indicated along *x*-axis) were calculated separately for genes that are also bound or not bound by E2f3 in NPCs. (**g**) Distribution of phylogenetic conservation scores of E2f3- and/or Ctcf-bound gene promoters, compared with all surveyed loci. Box plots demonstrating increased phylogenetic conservation of promoters bound by both E2f3 and Ctcf compared with those bound by a single factor. Asterisks, $P < 0.05$ by Wilcoxon's rank-sum test. (**h**) Number of enhancer–promoter units (EPUs) for E2f3- and/or Ctcf-bound promoters. The average number of enhancers, as determined by ref. 29 in embryonic brain, predicted to engage in enhancer–promoter units is represented for various categories of promoters identified in our experiments
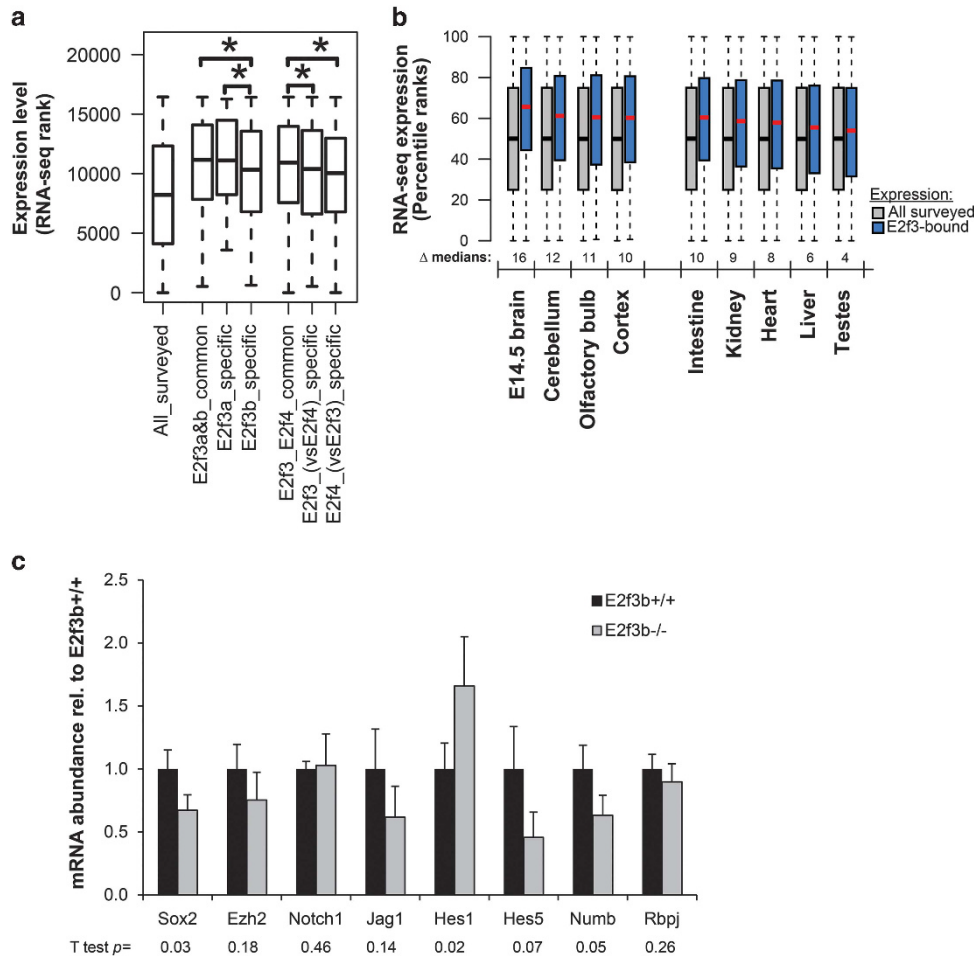
Figure 5 In NPCs, E2fs bind the TSS of genes that are transcriptionally active in neurogenic tissues. (**a**) Box plots display ranked expression data for genes belonging to distinct classes of E2f3 and E2f4 targets (gene classes indicated along the *x*-axis). Asterisks, $P < 0.05$ by Wilcoxon's rank-sum test. In addition, all the E2f-bound categories are significantly more highly expressed than the group of surveyed genes. (**b**) Expression levels of E2f3-bound and 'all surveyed' genes as measured by RNA-seq in various tissues. To evaluate if E2f3-bound genes are differently expressed than the group of surveyed genes, the difference between the median percentile rank of these two groups is given for each tissue (Δ medians). (**c**) Gene deregulation in E2f3-deficient NPCs. mRNA from wild-type and E2f3b $-/-$ neurospheres was quantified using real-time RT-PCR. Data are normalized using Gapdh as invariant control, and are reported as fold over the signal detected in wild-type samples. $n = 5$ wild-type and 6 mutant embryos from two age-matched litters derived from the same father. The results of unpaired two-tailed *t*-tests are shown below each gene. Error bars indicated S.E.M.

in fate determination, and E2f3 is associated with specialized, tissue-specific differentiation programs.

We found that E2fs are poised to control diverse NPC fate choices, including self-renewal, precursor maintenance, cell death and neuron generation, by directly binding the promoters of hundreds of cell fate-associated genes. The majority of target genes we identified are bound by multiple E2f family members, all predominantly associated with gene activation in NPCs, suggesting extensive functional redundancy by E2fs across their target genes. This finding underscores the importance of identifying factor-specific targets in an unbiased manner to fully appreciate the genetic mechanisms driving biological phenomena.

Our discovery that E2f3&4 are enriched at the promoters of such a large number of cell fate genes suggests that fate control by the pRb/E2f pathway may be largely independent from cell cycle regulation. Indeed, we have reported that the reduced/enhanced neurogenesis caused by loss of E2f3a/ E2f3b occurs without any changes to the length of cell cycle

phases in affected NPCs.[8] Another recent study demonstrated that pRb inhibits the ability of murine fibroblasts to reprogram to a pluripotent state by inhibiting E2f transcriptional activity, and that this process is also unaccompanied by changes in cell cycle dynamics.[44] Thus it appears that, at least in cells that are actively making stem cell fate decisions, pRb/E2f factors can direct these decisions independently from cell cycle regulation. Our data suggest that this is through direct regulation of networks of cell fate-associated genes.

We unexpectedly identified Ctcf as a potential regulatory co-factor for E2f3 at differentiation genes in NPCs. While Ctcf is a well-known insulator protein in mammals,[45] genome-wide studies have recently demonstrated that 20–25% of Ctcf-bound sites are located within proximal promoter regions,[46] and enrichment of E2f binding motifs has been identified among Ctcf-bound regions in multiple cell types.[47,48] Furthermore, the chromatin regulatory function of Ctcf has recently been established as a tumor suppressive mechanism,[49] akin to pRb/E2f. These studies substantiate our discovery, and
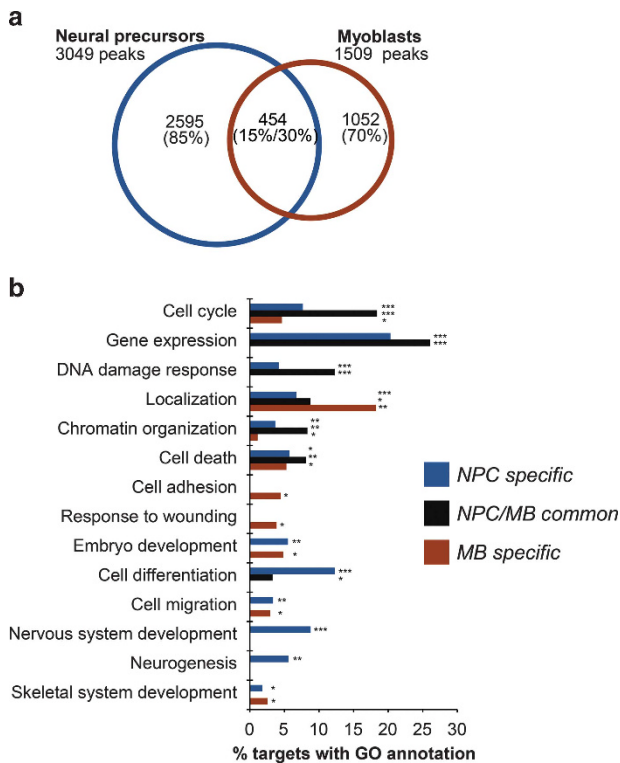
a

Neural precursors
3049 peaks

Myoblasts
1509 peaks

2595
(85%)

454
(15%/30%)

1052
(70%)

b

NPC specific

NPC/MB common

MB specific

% targets with GO annotation

Figure 6 E2f3 binds predominantly unique target genes in neural versus muscle precursors. (**a**) The number and overlap of peaks identified for E2f3, using a common pan-E2f3 antibody, in neural precursors (NPC) and myoblasts (MB) are shown. (**b**) GO analysis of NPC- and MB-common and specific genes. The percentage of genes from each category with specific functional annotations (indicated along *the y*-axis) is graphed. *P*-values are also indicated (*$P < 1.0E - 4$, **$P < 1.0E - 10$, ***$P < 1.0E - 20$). E2f3-specific peaks in NPCs are uniquely enriched in processes related to differentiation and development

suggest the possibility that the E2f3–Ctcf connection we have uncovered may be relevant in other cell types. However, we foresee a certain degree of tissue specificity in this connection, since it seems inexistent in muscle cells. More functional analyses are required to determine if this association is unique to NPCs or is more widespread, and to clarify how Ctcf–E2f interactions impact transcription.

We also discovered that E2f3 is more strongly associated with differentiation-related genes than is E2f4. This observation is supported biologically by the fact that NPCs deficient in E2f3 but not E2f4 exhibit overt neuronal differentiation defects.[8,10] We were also surprised to observe that E2f3 binds target sites in a highly tissue-specific manner in NPCs compared with MBs. Genetic mouse models have highlighted E2f3 as an important regulator of differentiation in a number of tissues, including brain, skeletal and cardiac muscle, cartilage and adipocytes,[8,24,50–53] and our data now suggest that regulation of E2f3-specific target genes is likely to contribute to distinct differentiation programs in these and potentially other cell types. To our knowledge, ours is the first study to demonstrate such a substantial degree of tissue specificity among genomic binding sites for any E2f factor. This phenomenon does not appear to be the result of differential isoform usage by E2f3 in NPCs versus MBs, as the common antibody used in both studies pulled down binding sites of both

E2f3a&b, with E2f3b-bound sites making up the majority in both cell types (our data and ref. 24). It will be important for future studies to examine E2f3a&b binding patterns in other tissue and cell types to reveal the extent by which E2f3 may function as a determinant of tissue-specific differentiation programs.

In conclusion, we have provided the first comprehensive, genome-wide view of the genes and associated cellular functions targeted by the pRb/E2f pathway in neural cells, and have thereby significantly enriched our understanding of how the cell cycle machinery impacts cell fate control and tissue-specific differentiation.

Materials and Methods

Mouse models and cell culture. Germline E2f3a- and E2f3b-deficient mice were generated originally by GL and were maintained on an FVB/N background.[8,54] All cells used in this study were cultured as neurospheres. Neural precursors were obtained by dissection of the ganglionic eminence from developing embryos at gestational age E14.5, and cells were cultured as neurospheres as previously described.[11] E2f3a $- / -$ and E2f3b $- / -$ NPCs were obtained from crosses of either two heterozygous or two homozygous (knock-out) animals. Animal experiments were approved by the University of Ottawa's Animal Care Committee, which abides by the guidelines of the Canadian Council on Animal Care. Mice were anesthetized using pentobarbital and killed by cervical dislocation.

Chromatin immunoprecipitation. ChIP assays were performed and quantified as previously described,[8] using 2 μg of specific antibody and 20 μg of chromatin for each reaction. ChIP data from WT cells are from a combination of E2f3a+/+ and E2f3b+/+ experiments; we observed no obvious discrepancies in levels of chromatin enrichment with any primers pair assessed in WT cells from either colony. All chromatin preparations were made from animals between the ages of 2- and 4-months old. Statistical analysis of ChIP data was performed using an unpaired two-tailed t-test, with differences considered significant with a *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

ChIP antibodies and primers. Identical E2f and IgG antibodies were used for ChIP–chip and conventional ChIP experiments. These include rabbit-specific antibodies against E2f3 (sc-878) and E2f4 (sc-866), and normal rabbit IgG as a control antibody for ChIP (all from Santa-Cruz, Dallas, TX, USA). The primer pairs used to amplify enriched chromatin at gene promoters are given in Supplementary Table S8.

Quantitative RT-PCR. Total RNA was isolated from neurospheres derived from WT and E2f3b $- / -$ embryos using RNeasy columns (Qiagen, Toronto, ON, Canada) and DNase treatment (Invitrogen, Burlington, ON, Canada). The embryos were obtained from two age-matched litters from the same father, from E2f3b-heterozygous crosses. Reverse-transcription followed by real-type PCR quantitation with the SYBR Green method were performed as previously described[55] using gene-specific primers spanning introns, whenever possible. Expression was normalized to the Gapdh housekeeping gene.

Electrophoretic mobility shift assay. Electrophoretic mobility shift assay was performed as previously described.[9] The material was neurospheres at stage E14.5 cultured from either littermate WT and E2f3 $- / -$ embryos, or littermate WT and E2f3a $- / -$ embryos. The antibodies used were sc-878, to detect E2f3, and sc-193 and sc-633 to detect E2f1 and E2f2.

ChIP–chip and assignment of E2f target genes. A custom microarray manufactured by Agilent and containing ~ 1 million probes was used to survey the proximal promoters of 24 654 mouse transcripts. Details are provided in Supplementary Text S1. Chromatin from ~ 50 million cells was fragmented and immunoprecipitated with the anti-E2F3 antibody. The immunopurified and input DNA (pre-IP) samples were amplified by LM-PCR, fluorescently labeled and co-hybridized on a DNA microarray custom made by Agilent Technologies (Mississauga, ON, Canada), essentially as described previously.[55] Experimental details and data analysis methods are provided in Supplementary Text S1. The data have been deposited on the NCBI Gene Expression Omnibus under the accession number GSE54941. The

complete set of E2f3-bound regions is given in Supplementary Table S1. Mapping is done on the mm9 (NCBI build 37) mouse genome assembly.

Comparisons between ChIP–chip and ChIP-Seq data sets. The overlap between various groups of bound regions was determined using the UCSC genome table browser and the intersection tool, requiring an overlap of at least 1 bp. The E2f3-binding data reported by Asp *et al.*[24] in skeletal MBs were converted to mm9 coordinates using the liftOver tool. The regions bound by Ctcf and RNA pol II in embryonic brain used in our comparisons were obtained from the NCBI GEO, under accession GSE29218.[37] The regions bearing certain histone modification marks in neural precursors were taken from GSE12241.[56] Ctcf-bound regions in C2C12 MBs were those reported by the Wold lab (Caltech, University in Pasadena, Pasadena, CA, USA) as part of the ENCODE project. For most analyses of gene promoters, we defined proximal promoters as the region from 2 kb upstream to 2 kb downstream of the TSS. The promoter–enhancer connections were based on the 'enhancer–promoter units' predicted to exist in embryonic brain.[37] E2F3- or Ctcf-bound proximal promoters that overlap with the promoter component of the E14.5 brain enhancer–promoter units were identified, and the number of enhancers predicted to link to these promoters was counted.

GO analysis. The enrichment of specific ontologies among the genes bound by E2f3 was assessed using the program GREAT, with the 'single nearest gene' within 1 kb association rule. As 'background regions', we used the entire set of genomic loci that have been surveyed by our microarrays. Significance values used were the hypergeometric test with correction for multiple hypothesis testing via the Benjamini algorithm. In this case, differences are considered significant with the following P-values: $*P < 1.0E - 4$, $**P < 1.0E - 10$, $***P < 1.0E - 20$.

DNA sequence analyses. The E2f3-bound regions were analyzed for the presence, and enrichment, of certain transcription factor DNA-binding sequences using CisGenome[58] and the position weight matrices from TRANSFAC, at a cut-off of log likelihood ratio of 500. The frequency of sequence motifs in the transcription factor-bound loci was compared with that in control regions, a group of randomly selected sequenced from the set of sequences surveyed by the microarray. For the E2F motif, we used TRANSFAC matrix 738. We used the Ctcf binding matrix reported by Schmidt *et al.*[59] for murine Ctcf. The phylogenetic conservation (phastCons scores) of sequences was analyzed using CisGenome.

Gene expression profiling. Data from RNA-Seq performed in embryonic mouse brain at E14.5 were downloaded from NCBI GEO, under accession number GSM850906.[37] The fragments per kilobase of exon per million fragments mapped (FPKM) expression values of genes were assigned to individual E2f3 target sites using GREAT and the 'single nearest gene' association rule. To obtain a non-parametric view of gene expression from this data set, we analyzed genes based on their rank within the RNA-Seq data set: the entire list of detected genes was sorted in order of increasing fragments per kilobase of exon per million values, and genes were assigned a rank (lowest rank number meaning lowest expression level). Box plots were made using R[60]; they represent the median value (thick horizontal line), the first and third quartile boundaries (boxes) and 1.5 times the interquartile distance (whiskers). Statistical significance was assessed using the Wilcoxon rank-sum test.

We also determined the expression profile of E2f3 target genes discovered in our ChIP–chip analysis across a panel of over 90 mouse tissues and cell lines. Gene expression data for all E2f3 target genes in NPCs were obtained from the mouse Gene Atlas V3,[61] and was analyzed and normalized as previously described.[55] K-means clustering was performed on the gene expression data using the Cluster 3 program,[62] and was visualized using Java Treeview[63] to identify tissues and cell types that demonstrate similar gene expression profiles.

Conflict of Interest

The authors declare no conflict of interest.

Author contributions

LMJ, RSS and AB conceived and designed the experiments. LMJ performed most of the experiments, and LMJ and AB performed the data analyses and wrote the manuscript. YL, CAP and VR provided technical training and assistance with ChIP experiments, DD-T performed Western blots, and all provided thoughtful discussion towards manuscript preparation. YL further assisted with training for the ChIP-on-chip studies, data analysis and preparation of data files for publication, and performed the quantitative RT-PCR measurements. GL, S-YT and SB generated and provided the mouse models used in this study. All authors read and provided feedback on the manuscript, and approved the final version.

1. Arai Y, Pulvers JN, Haffner C, Schilling B, Nüsslein I, Calegari F *et al.* Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat Commun* 2011; **2**: 154.
2. Calegari F, Haubensak W, Haffner C, Huttner WB. Selective lengthening of the cell cycle in the neurogenic subpopulation of neural progenitor cells during mouse brain development. *J Neurosci* 2005; **25**: 6533–6538.
3. Lange C, Huttner WB, Calegari F. Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Stem Cell* 2009; **5**: 320–331.
4. Lim S, Kaldis P. Loss of Cdk2 and Cdk4 induces a switch from proliferation to differentiation in neural. *Stem Cells* 2012; **30**: 1509–1520.
5. Chen D, Opavsky R, Pacal M, Tanimoto N, Wenzel P, Seeliger MW *et al.* Rb-mediated neuronal differentiation through cell-cycle-independent regulation of E2f3a. *PLoS Biol* 2007; **5**: e179.
6. Cooperkuhn C, Vroemen M, Brown J, Ye H, Thompson MA, Winkler J *et al.* Impaired adult neurogenesis in mice lacking the transcription factor E2F1. *Mol Cell Neurosci* 2002; **21**: 312–323.
7. Ferguson KL, McClellan KA, Vanderluit JL, McIntosh WC, Schuurmans C, Polleux F *et al.* A cell-autonomous requirement for the cell cycle regulatory protein, Rb, in neuronal migration. *EMBO J* 2005; **24**: 4381–4391.
8. Julian LM, Vandenbosch R, Pakenham CA, Andrusiak MG, Nguyen AP, McClellan KA *et al.* Opposing regulation of Sox2 by cell-cycle effectors E2f3a and E2f3b in neural stem cells. *Stem Cell* 2013; **12**: 440–452.
9. Mcclellan KA, Ruzhynsky VA, Douda DN, Vanderluit JL, Ferguson KL, Chen D *et al.* Unique requirement for Rb/E2F3 in neuronal migration: evidence for cell cycle-independent functions. *Mol Cell Biol* 2007; **27**: 4825–4843.
10. Ruzhynsky VA, McClellan KA, Vanderluit JL, Jeong Y, Furimsky M, Park DS *et al.* Cell cycle regulator E2F4 is essential for the development of the ventral telencephalon. *J Neurosci* 2007; **27**: 5926–5935.
11. Vanderluit JL, Ferguson KL, Nikoletopoulou V, Parker M, Ruzhynsky V, Alexson T *et al.* p107 regulates neural precursor cells in the mammalian brain. *J Cell Biol* 2004; **166**: 853–863.
12. Kippin TE. p21 loss compromises the relative quiescence of forebrain stem cell proliferation leading to exhaustion of their proliferation capacity. *Genes Dev* 2005; **19**: 756–767.
13. Nishino J, Kim I, Chada K, Morrison SJ. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf expression. *Cell* 2008; **135**: 227–239.
14. Vanderluit JL, Wylie CA, McClellan KA, Ghanem N, Fortin A, Callaghan S *et al.* The Retinoblastoma family member p107 regulates the rate of progenitor commitment to a neuronal fate. *J Cell Biol* 2007; **178**: 129–139.
15. Ferguson KL, Vanderluit JL, Hébert JM, McIntosh WC, Tibbo E, MacLaurin JG *et al.* Telencephalon-specific Rb knockouts reveal enhanced neurogenesis, survival and abnormal cortical development. *EMBO J* 2002; **21**: 3337–3346.
16. Svoboda DS, Paquin A, Park DS, Slack RS. Pocket proteins pRb and p107 are required for cortical lamination independent of apoptosis. *Dev Biol* 2013; **384**: 101–113.
17. Mcclellan KA, Vanderluit JL, Julian LM, Andrusiak MG, Dugal-Tessier D, Park DS *et al.* The p107/E2F pathway regulates fibroblast growth factor 2 responsiveness in neural precursor cells. *Mol Cell Biol* 2009; **29**: 4701–4713.
18. Andrusiak MG, McClellan KA, Dugal-Tessier D, Julian LM, Rodrigues SP, Park DS *et al.* Rb/E2F regulates expression of neogenin during neuronal migration. *Mol Cell Biol* 2011; **31**: 238–247.
19. Ghanem N, Andrusiak MG, Svoboda D, Al Lafi SM, Julian LM, McClellan KA *et al.* The Rb/E2F pathway modulates neurogenesis through direct regulation of the Dlx1/Dlx2 bigene cluster. *J Neurosci* 2012; **32**: 8219–8230.
20. Xu X, Bieda M, Jin VX, Rabinovich A, Oberley MJ, Green R *et al.* A comprehensive ChIP chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members. *Genome Res* 2007; **17**: 1550–1561.
21. Chen H-Z, Tsai S-Y, Leone G. RB and E2f: Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer* 2009; **9**: 785–797.

22. Conboy CM, Spyrou C, Thorne NP, Wade EJ, Barbosa-Morais NL, Wilson MD *et al*. Cell cycle genes are the evolutionarily conserved targets of the E2F4 transcription factor. *PLoS One* 2007; **2**: e1061.

23. Lee BK, Bhinge AA, Iyer VR. Wide-ranging functions of E2F4 in transcriptional activation and repression revealed by genome-wide analysis. *Nucleic Acids Res* 2011; **39**: 3558–3573.

24. Asp P, Acosta-Alvear D, Tsikitis M, van Oevelen C, Dynlacht BD. E2f3b plays an essential role in myogenic differentiation through isoform-specific gene regulation. *Genes Dev* 2009; **23**: 37–53.

25. von Eyss B, Maaskola J, Memczak S, Möllmann K, Schuetz A, Loddenkemper C *et al*. The SNF2-like helicase HELLS mediates E2F3-dependent transcription and cellular transformation. *EMBO J* 2011; **31**: 972–985.

26. Ren B. E2F integrates cell cycle progression with DNA repair, replication, and G2/M checkpoints. *Genes Dev* 2002; **16**: 245–256.

27. Attwool C, Lazzerini DE, Helin K. The E2F family: specific functions and overlapping interests. *EMBO J* 2004; **23**: 4709–4716.

28. DeGregori J, Johnson DG. Distinct and overlapping roles for E2F family members in transcription, proliferation and apoptosis. *Curr Mol Med* 2006; **6**: 739–748.

29. Dimova DK, Dyson NJ. The E2F transcriptional network: old acquaintances with new faces. *Oncogene* 2007; **24**: 2810–2826.

30. Iaquinta PJ, Lees JA. Life and death decisions by the E2F transcription factors. *Curr Opin Cell Biol* 2007; **19**: 649–657.

31. McClellan KA, Slack RS. Specific in vivo roles for E2Fs in differentiation and development. *Cell Cycle* 2007; **6**: 2917–2927.

32. Humbert PO, Verona R, Trimarchi JM, Rogers C, Dandapani S, Lees JA *et al*. . E2f3 is critical for normal cellular proliferation. *Genes Dev* 2000; **14**: 690–703.

33. Muller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E *et al*. E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev* 2001; **15**: 267–285.

34. Wang A, Schneider-Broussard R, Kumar AP, MacLeod MC, Johnson DG. Regulation of BRCAI Expression by the Rb-E2F Pathway. *J Biol Chem* 2000; **275**: 4532–4536.

35. Marqués-Torrejón MÁ, Porlan E, Banito A, Gómez-Ibarlucea E, Lopez-Contreras AJ, Fernández-Capetillo O *et al*. Cyclin-dependent kinase inhibitor p21 controls adult neural stem cell expansion by regulating Sox2 gene expression. *Stem Cell* 2013; **12**: 88–100.

36. Elkon R. Genome-wide in silico identification of transcriptional regulators controlling the cell cycle in human cells. *Genome Res* 2003; **13**: 773–780.

37. Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S *et al*. A map of the cis-regulatory sequences in the mouse genome. *Nature* 2012; **488**: 116–120.

38. Adams MR, Sears R, Nuckolls F, Leone G, Nevins JR. Complex transcriptional regulatory mechanisms control expression of the E2F3 locus. *Mol Cell Biol* 2000; **20**: 3633–3639.

39. Leone G, Nuckolls F, Ishida S, Adams M, Sears R, Jakoi L *et al*. Identification of a novel E2F3 product suggests a mechanism for determining specificity of repression by Rb proteins. *Mol Cell Biol* 2000; **20**: 3626–3632.

40. He Y, Armanious MK, Thomas MJ, Cress WD. Identification of E2F-3B, an alternative form of E2F-3 lacking a conserved N-terminal region. *Oncogene* 2000; **19**: 3422–3433.

41. Herrup K. Post-mitotic role of the cell cycle machinery. *Current Opin Cell Biol* 2013; **25**: 711–716.

42. Ruzhynsky V.A, McClellan KA, Vanderluit JL, Jeong Y, Furimsky M, Park DS *et al*. Cell cycle regulator E2F4 is essential for the development of the ventral telencephalon. *J Neurosci* 2007; **27**: 5926–5935.

43. Chen D, Chen Y, Forrest D, Bremner R. E2f2 induces cone photoreceptor apoptosis independent of E2f1 and E2f3. *Cell Death Differ* 2013; **20**: 931–940.

44. Kareta MS, Gorges LL, Hafeez S, Benayoun BA, Marro S, Zmoos A *et al*. . Inhibition of pluripotency networks by the Rb tumor suppressor restricts reprogramming and tumorigenesis. *Cell Stem Cell* 2015; **16**: 39–50.

45. Phillips-Cremins JE, Corces VG. Chromatin insulators: linking genome organization to cellular function. *Mol Cell* 2013; **50**: 461–474.

46. Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD. Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* 2007; **8**: 427–437.

47. Martin D, Pantoja C, Fernández Miñán A, Valdes-Quezada C, Moltó E, Matesanz F *et al*. Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat Struct Mol Biol* 2011; **18**: 708–714.

48. Whitfield TW, Wang J, Collins PJ, Partridge EC, Aldred SF, Trinklein ND *et al*. Functional analysis of transcription factor binding sites in human promoters. *Genome Biol* 2012; **13**: R50.

49. Kemp CJ, Moore JM, Moser R, Bernard B, Teater M, Smith LE *et al*. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep* 2014; **7**: 1020–1029.

50. King JC, Moskowitz IP, Burgon PG, Ahmad F, Stone JR, Seidman JG *et al*. E2F3 plays an essential role in cardiac development and function. *Cell Cycle* 2008; **7**: 3775–3780.

51. Chong J-L, Tsai SY, Sharma N, Opavsky R, Price R, Wu L *et al*. E2f3a and E2f3b contribute to the control of cell proliferation and mouse development. *Mol Cell Biol* 2009; **29**: 414–424.

52. Danielian PS, Friesenhahn LB, Faust AM, West JC, Caron AM, Bronson RT *et al*. E2f3a and E2f3b make overlapping but different contributions to total E2f3 activity. *Oncogene* 2008; **27**: 6561–6570.

53. Chen L, Cui J, Hou J, Long J, Li C, Liu L *et al*. A novel negative regulator of adipogenesis: microRNA-363. *Stem Cells* 2014; **32**: 510–520.

54. Tsai S-Y, Opavsky R, Sharma N, Wu L, Naidu S, Nolan E *et al*. Mouse development with a single E2F activator. *Nature* 2008; **454**: 1137–1141.

55. Liu Y, Chu A, Chakroun I, Islam U, Blais A. Cooperation between myogenic regulatory factors and SIX family transcription factors is important for myoblast differentiation. *Nucleic Acids Res* 2010; **38**: 6857–6871.

56. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G *et al*. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; **448**: 553–560.

57. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB *et al*. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010; **28**: 495–501.

58. Ji H, Wong WH. TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics* 2005; **21**: 3629–3636.

59. Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C *et al*. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 2012; **148**: 335–348.

60. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2013. http://www.R-project.org/.

61. Lattin JE, Schroder K, Su AI, Walker JR, Zhang J, Wiltshire T *et al*. Expression analysis of G protein-coupled receptors in mouse macrophages. *Immunome Res* 2008; **4**: 5.

62. de Hoon MJL, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* 2004; **20**: 1453–1454.

63. Saldanha AJ. Java Treeview—extensible visualization of microarray data. *Bioinformatics* 2004; **20**: 3246–3248.