A unique bipartite Polycomb signature regulates stimulus-response transcription during development

Taro Kitazawa1,6, Dania Machlab2,3,6, Onkar Joshi1, Nicola Maiorano1, Hubertus Kohler1, Sebastien Ducret1, Sandra Kessler1, Henrik Gezelius4,5, Charlotte Soneson1,2, Panagiotis Papasaikas1,2, Guillermínía López-Bendito4, Michael B. Stadler1,2 and Filippo M. Rijli1,3✉

Rapid cellular responses to environmental stimuli are fundamental for development and maturation. Immediate early genes can be transcriptionally induced within minutes in response to a variety of signals. How their induction levels are regulated and their untimely activation by spurious signals prevented during development is poorly understood. We found that in developing sensory neurons, before perinatal sensory-activity-dependent induction, immediate early genes are embedded into a unique bipartite Polycomb chromatin signature, carrying active H3K27ac on promoters but repressive Ezh2-dependent H3K27me3 on gene bodies. This bipartite signature is widely present in developing cell types, including embryonic stem cells. Polycomb marking of gene bodies inhibits mRNA elongation, dampening productive transcription, while still allowing for fast stimulus-dependent mark removal and bipartite gene induction. We reveal a developmental epigenetic mechanism regulating the rapidity and amplitude of the transcriptional response to relevant stimuli, while preventing inappropriate activation of stimulus-response genes.

During development, cells are exposed to a variety of distinct environmental signals to which they may need to rapidly respond in a spatiotemporally regulated manner, in order to keep their differentiation schedule. Stimulus-response genes are essential for rapid cellular responses to extracellular signals1–3. Among them, immediate early genes (IEGs) are induced in multiple cell types within minutes in a stimulus-dependent manner, often encoding transcription factors (for example, Fos and Egr1), which in turn regulate the expression of downstream late-response genes (LRGs) through activation of enhancers4–6. Before induction, IEGs share key regulatory properties, which poise them for rapid stimulus-dependent activation. In general, these include accessible promoters and enhancers bound by serum response factor, nuclear factor-κB, cyclic AMP response element-binding protein (CREB) and/or activator protein-1 transcription factors, which are posttranslationally modified upon stimulus response, as well as transcriptionally permissive histone modifications (H3K4me2/3) and paused RNA polymerase II (RNAPII)7,8. Despite their shared organization, differences in transcription initiation, elongation or mRNA processing and stability may result in IEG induction differences7,9. Moreover, IEGs are both general (that is, the same IEGs are induced in most cell types in response to different stimuli) and cell-type specific (responding to specific signals in different cell types)1,8,9–12. How spatiotemporal regulation and specificity of the IEG transcriptional response is achieved in developing cells, and how untimely induction of IEGs in response to spurious signals is prevented are poorly understood.

Here, we asked whether and how chromatin states might also contribute to stimulus-dependent transcriptional regulation of IEGs during development, choosing the mouse developing somatosensory neurons as a suitable model. We then further confirmed the general validity of our findings in developing neural crest, heart, liver and embryonic stem cells (ESCs). We discovered, and functionally investigated, a unique H3K27ac/H3K27me3 bipartite chromatin signature, which provides an epigenetic mechanism to modulate the rapidity and amplitude of the transcriptional response of inducible IEGs to distinct stimuli during development. Our findings support the involvement of Polycomb (Pc)-dependent H3K27me3 on the gene body in inhibiting the productive elongation of RNAPII on bipartite genes. While strong stimuli allow for the rapid removal of Pc marking of gene bodies and fast transcriptional induction, Pc marking of gene bodies of bipartite stimulus-response genes may establish a threshold to prevent rapid transcriptional induction of IEGs in response to suboptimal and/or nonphysiologically relevant levels of environmental stimuli.

Results

Transcriptional and chromatin profiling of activity-regulated genes in developing neurons. During early postnatal sensory neuron development, IEGs and LRGs are transcriptionally induced by sensory experience, which drives neuronal and circuit maturation13,14. In the mouse somatosensory system, topographic representations of the mystacial vibrissae (whiskers) on the face are generated at brainstem, thalamus and cortical levels13,14. In the brainstem, the whisker-related neuronal modules, or barrelettes, are generated in the ventral principal trigeminal sensory nucleus (vPrV), and sensory neuronal activity is required at perinatal/early postnatal stages for the maturation of barrelette neuron connectivity and map formation13,14.

To characterize IEG and LRG activity-response genes (ARGs) in developing barrelette neurons, we set out a genetic strategy to isolate

1Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. 2Swiss Institute of Bioinformatics, Basel, Switzerland. 3University of Basel, Basel, Switzerland. 4Instituto de Neurociencias de Alicante, Universidad Miguel Hernández-Consejo Superior de Investigaciones Científicas (UMH-CSIC), Sant Joan d’Alacant, Spain. 5Present address: Science for Life Laboratory, Solna, Sweden. 6These authors contributed equally: Taro Kitazawa, Dania Machlab. ✉ e-mail: filippo.rijli@fmi.ch
embryonic day (E) 10.5 mitotic progenitors and postmitotic barrelette neurons at E14.5 (early postmitotic), E18.5 (perinatal) and postnatal day (P) 4 (consolidated barrelette stage) by FACS, and we profiled them by mRNA sequencing (mRNA-seq; Smart-seq2) and chromatin immunoprecipitation followed by sequencing (ChiP–seq; ChiPmentation) of the Pc-dependent repressive H3K27me3 and active H3K4me2 and H3K27ac histone modifications, and chromatin accessibility by an assay for transposable-accessible chromatin followed by sequencing (ATAC–seq; Methods, Fig. 1a, Extended Data Fig. 1a–e, Supplementary Table 1, Supplementary Figs. 1 and 2 and Supplementary Note).

To identify ARGs induced in barrelette neurons at the beginning of the sensory-dependent maturation period (E18.5–P2/P3)19, we collected E18.5 Kir2.1-overexpressing, activity-deprived, vPrV postmitotic barrelette neurons by FACS sorting (Extended Data Fig. 1c–p, Supplementary Table 1, Supplementary Fig. 2, Supplementary Note and Methods), profiled them by mRNA-seq (Smart-seq2), and compared to E14.5 and E18.5 vPrV wild-type barrelette neurons. Among the genes with undetectable or low basal expression level (reads per kilobase per million mapped reads (RPKM) <3) in E14.5 barrelette neurons, we identified 56 genes, referred to as barrelette sensory ARGs (bsARGs; Supplementary Table 2), which were upregulated at E18.5 in a neuronal activity-dependent manner (Extended Data Fig. 1q–s, Supplementary Note and Methods). Barrelette sensory ARGs comprised 4 IEGs, namely Fos, Egr1, Junb and Zfp36 (Fig. 1b and Supplementary Table 2), and at least 25 putative LRGs (for example, Cldn8 and Osmr, Supplementary Table 3). We next identified additional ARGs referred to as non-barrelette ARGs (nbARGs; n=83; Methods and Supplementary Note), which included both LRGs and 12 IEGs and were transcriptionally induced by distinct activity-dependent stimuli in neuronal types other than barrelette neurons20, but that displayed undetectable or low basal expression level (RPKM <3) in E14.5, E18.5 and P4 barrelette neurons.

PC group proteins regulate dynamics and plasticity of gene expression during development13–22. We found that in E14.5 barrelette neurons, 32/56 (57%) and 67/83 (84%) of bsARGs and nbARGs, respectively, were embedded in H3K27me3 domains of Pc-repressive chromatin (Methods) with, however, H3K4me2/AIDS signatures (Fig. 1c and Extended Data Fig. 1t,u). In analysis of their chromatin profile, only 4 of 16 (25%) IEGs (Junb, Egr3, Egr4 and Atf3) displayed a conventional Pc bivalent signature (Fig. 1d), that is, with promoters marked by both active H3K4me2 and repressive H3K27me3 histone modifications. Interestingly, 12 of 16 (75%) IEGs displayed a unique distinct ‘bipartite’ Pc signature (Fig. 1d), that is, with promoters marked by both active H3K4me2 and repressive H3K27me3 histone modifications. Immediately, 12 of 16 (75%) IEGs displayed a unique distinct ‘bipartite’ Pc signature (Fig. 1d); see genome browser snapshots at Fos, Egr1, Fosb and Nr4a3 in Fig. 1e; Extended Data Fig. 2a). Namely, H3K27me3 deposition was restricted to their gene bodies, whereas the accessible H3K4me2+ promoters were devoid of H3K27me3 and instead carried the active mark H3K27ac, notably with no or only low basal levels of detected mRNA. H3K27me3 on gene bodies did not stretch further than 2–3kb downstream of the transcription start site (TSS), even when the gene was longer (for example, Nr4a3; Extended Data Fig. 2a). H3K27ac deposition at promoters of bipartite IEGs was not induced by the dissociation procedure (Extended Data Fig. 2b and Methods). Conversely, we found that among the remaining 83 of 99 H3K4me2+/H3K27me3+/ATAC+ ARGs, which included putative barrelette neuron LRGs and non-barrelette neuron LRGs16–18 (for example, Osmr and Pdlim1, respectively; Fig. 1e), 66 of 83 (80%) were in a bivalent state, whereas only 17 of 83 (20%) carried the bipartite Pc signature (Fig. 1d and Methods).

In summary, at prenatal stages, the rapidly inducible IEGs are preferentially in a Pc bipartite state, while the LRGs are preferentially enriched with a Pc bivalent signature (Fig. 1d,e). Similarly to developing barrelette neurons, the Pc bipartite signature was also present at IEGs in prenatal cortical progenitors and postmitotic neurons, although neither in adult excitatory neurons nor in 7-d cultured embryonic cortical neurons (Extended Data Fig. 2a). Thus, the bipartite chromatin organization is specifically established at IEGs during prenatal neuronal development.

The bipartite signature is found on stimulus-response genes during development and is not restricted to neurons. We next investigated the genome-wide distribution of the Pc bipartite chromatin signature. We assigned each gene with a ‘bipartiteness’ score related to their promoter H3K27ac and gene body H3K27me3 levels and a ‘bivalency’ score related to H3K27me3 and H3K4me2 at promoters (Methods and Extended Data Fig. 3a,b). Considering the estimated false-positive rates of this scoring approach, we conservatively evaluated the total numbers of true bipartite genes from at least 140 at E10.5, to 177 at E14.5, to 219 at E18.5 and decreasing to 113 at P4 in barrelette neurons (Fig. 2a and Methods). At all stages, approximately 1,500 genes were instead in a bivalent state (Fig. 2a and Methods). Aggregate profile plots of chromatin marks of the top 100 E14.5 barrelette neuron genes scored as bipartite (E14.5Bip) or bivalent (E14.5Biv; Methods) further confirmed their clearly distinct chromatin signatures (Fig. 2b and Extended Data Fig. 3c).

In addition to IEGs, Gene Ontology analysis of E14.5Bip genes identified genes encoding transcriptional regulators and transmembrane domain receptors responding to distinct signaling pathways including bone morphogenetic protein and transforming growth factor-β signaling, voltage-gated ion channels, and dendritic, axonal and synaptic genes (Fig. 2c and Supplementary Table 4).

Furthermore, by our ranking method, we additionally found 124, 99, 185 and 107 genes carrying the bipartite chromatin signature in mouse E14.5 heart tissue, E14.5 liver tissue, E10.5 cranial neural-crest-derived cells (NCCs) and ESCs, respectively (Fig. 2a–c, Extended Data Fig. 3d and Supplementary Table 4). Bipartite genes are tissue- and stage-specific as only a few bipartite genes were shared among the different cell types (Fig. 2d), including typical IEGs (for example, Fos, Jun, Fos2, Myb, Egr2 and Arc; Fig. 2c). Nonetheless, bipartite genes appear to be consistently 5–15% of the bivalent genes at all times and in all the distinct cell types analyzed (Fig. 2a).
Lastly, sequential ChIP–seq on E14.5 bulk hindbrain tissue and single-cell mRNA-seq (scRNA-seq; 10x Genomics) analysis of FACS-isolated E14.5 postmitotic barrelette neurons and E10.5 progenitors demonstrated that the H3K27ac and H3K27me3 histone marks coexist at the promoter and gene body of bipartite genes, correlating with low or undetectable mRNA expressions.
transcription (Supplementary Note, Fig. 2e,f and Extended Data Figs. 3e and 4a–e).

These results show that the bipartite signature is not an exclusive feature of developing neurons but is widely used during development, raising the intriguing possibility that it could regulate rapid IEG transcriptional inducibility.

The bipartite signature originates from bivalent chromatin and is dynamic during development. To investigate how the bipartite signature is established, maintained and resolved during development, we created a two-dimensional (2D) projection of autosomal genes according to chromatin accessibility, H3K27me3, H3K4me2 and H3K27ac levels at promoters and gene bodies (Extended Data Fig. 5a) using t-distributed stochastic neighbor embedding (t-SNE; Fig. 3a–d, Extended Data Fig. 5b–l, Supplementary Note and Methods). We generated a single map for E10.5 progenitors and a combined E14.5, E18.5 and P4 t-SNE map of chromatin states for postmitotic barrelette neurons (Fig. 3a,b, Extended Data Fig. 5b,c and Methods). Genes with similar chromatin patterns were grouped together, which also correlated with mRNA-seq data (Extended Data Fig. 5c).

Top-scoring bipartite and bivalent genes at E10.5 and postmitotic stages mapped to distinct, largely nonoverlapping, regions on the respective t-SNE maps (Fig. 3a–d, Extended Data Fig. 5f,g,i–l, Supplementary Note and Methods). Furthermore, genes mapping to the same region of the combined E14.5/E18.5/P4 t-SNE map revealed a stable chromatin state, unlike genes changing their localization between developmental stages (Fig. 3a–d, Extended Data Fig. 5h and Supplementary Note).

At P4, distinct fractions of the E14.5Bip genes had transitioned into productive transcription (Bip → Exp; RPKM > 3), bivalency (Bip → Biv) or remained bipartite (Bip → Bip; Extended Data Fig. 6a). As compared to E14.5, Bip → Exp genes displayed higher levels of H3K27ac, increased accessibility (ATAC-seq) and mRNA levels, and decreased H3K27me3, in contrast to genes that remained bipartite (Bip → Bip) or became bivalent (Bip → Biv; Extended Data Fig. 6a). The developmental progression through distinct bipartite patterns and into the active chromatin state of E14.5Bip genes could also be readily visualized as relocation of their position on the E10.5, E14.5 and P4 t-SNE plots (Fig. 3c); representative examples include FoxO1 and Bcl6 (involved in postmitotic neuronal fate through repression of Wnt/Notch/Fgf/Shh), Nr3c1 (glucocorticoid receptor) and Pkhdh3 (signal transduction in axon growth; Fig. 2c), while Figure 3d shows the fraction of E14.5Bip genes that switched to bivalency at P4 (Bip → Biv). Genome browser views of FoxO1 and Egr1 (Bip → Exp) and Gpr88 (Bip → Biv) confirmed the transcriptional and epigenetic changes (Fig. 3e,f and Extended Data Fig. 6b).

Moreover, by using circular chromosome conformation capture, coupled to high-throughput sequencing (4C-seq), we found that the bipartite signature at the Fos locus allowed for reciprocal physical contacts between its active enhancers and promoter, irrespective of productive transcription (Supplementary Note, Fig. 3e and Extended Data Fig. 6c).

Next, we investigated the developmental origin of the bipartite signature. At E10.5, about 50% of E14.5Bip genes were already in a bipartite state, as they mapped within the green contour region of the E10.5 t-SNE plot; however, as much as 40% of E14.5Bip genes were in a bivalent state in E10.5 progenitors, as they were contained within the bivalent red contour region (Fig. 3b; see Fig. 3f for a representative example, Gpr88). While bipartite and bivalent genes had similar CpG content and distribution (Extended Data Fig. 3f), the E14.5Bip promoters were enriched in nuclear factor-kB-related and forkhead box-related factor binding motifs (Extended Data Fig. 3g and Methods).

Thus, the bipartite state originates from bivalent chromatin in early progenitors and, during postmitotic neuron development, displays bidirectional dynamics, reverting into a bivalent state for a subset of genes, or resolving into productive transcription.

RNA polymerase II transcripts of bipartite genes are not efficiently processed to productive mRNA. We next investigated additional chromatin features of bipartite genes (Fig. 4a and Supplementary Note).

Moreover, E14.5Bip genes displayed dramatically lower productive mRNA levels than E14.5 non-bipartite genes with Bip-matching promoter H3K27ac levels (E14.5AcP; Fig. 4a, mRNA). To investigate why active bipartite promoters did not drive higher levels of productive transcription, we determined ChIP enrichment for distinct phosphorylated forms of the main subunit of RNAPII. The RNAPII C-terminal domain changes its serine phosphorylation pattern as RNAPII progresses from initiation (S5P) through productive transcription and elongation (S7P and S2P)28–30. Transcriptionally productive and elongating RNAPII-S5P+S7P+S2P* is detected at active genes, whereas not productively elongating RNAPII-S5P+S7P+S2P*, also little or not recognized by the 8WG16 antibody, is detected at Pc-repressed bivalent genes29,31,32.

We found a unique pattern of RNAPII at E14.5Bip genes. Namely, 8WG16, RNAPII-S5P and RNAPII-S7P levels at E14.5Bip promoters were similar to those at E14.5AcP promoters, and higher than non-Bip genes with low, Bip-matching, levels of productive mRNA transcription (E14.5mRNAlow; Methods) and E14.5Bip promoters (Fig. 4a,b). In contrast, around the E14.5Bip transcription end sites (TESs; Methods), the levels of RNAPII-S2P and H3K36me3, a mark of productive mRNA elongation into gene bodies29, were significantly lower than those in E14.5AcP genes, although higher than those in E14.5Bip genes and comparable to those in E14.5mRNAlow genes (Fig. 4b). Genome browser views of E14.5Bip FoxO1 and Egr1 loci confirmed that both RNAPII-S5P and RNAPII-S7P paused at the promoter–proximal first exon regions, while RNAPII-S7P and...
RNAPII-S2P levels were barely detectable in the H3K27me3+ gene body regions (Extended Data Fig. 6d). This distribution is generally shared by E14.5Bip genes (Fig. 4). In addition, total RNA analysis (Ovation SoLo RNA-seq; Methods) showed that E14.5Bip nascent RNA transcripts were not efficiently processed to productive mRNA (Extended Data Fig. 7a,b).
In summary, mRNA processivity of E14.5Bip genes is intermediate between bivalent (E14.5Biv) genes and genes with comparable H3K27ac promoter levels (E14.5AcP). We also demonstrate that mRNA elongation through the gene bodies of E14.5Bip genes is maintained at a low rate, in line with E14.5RNA-low genes (Fig. 4a,b), having similar promoter H3K27ac and RNAPII-S5P levels to E14.5AcP genes.

Polycomb marking of bipartite genes on gene bodies inhibits productive mRNA processing. Little is known about a potential role of Polycomb marking of bipartite genes on gene bodies inhibiting productive spliced mRNA in Ezh2-cKO^{1,2-BP} mutant cells than controls (Extended Data Fig. 7c). Moreover, accumulation of reads at gene TSS proximal regions was reduced in mutant compared to wild-type cells (Fig. 5c and Extended Data Fig. 7d). Moreover, likely as a direct result of ectopic Fos induction, 85 activity-regulated Fos-binding enhancers (Methods), which normally became open only in postnatal barrelette neurons, gained precocious accessibility in E14.5 FACS-isolated Ezh2 homozygous mutant neurons from bulk hindbrain (Ezh2-cKO^{1,2-BP}; Supplementary Fig. 3, Supplementary Table 1 and Methods), suggesting incorrect precocious activation of an early postnatal Fos-driven enhancer program (Fig. 5d, Extended Data Fig. 8a-b and Supplementary Note).

Next, we investigated the levels and distribution of elongation marks in Ezh2 mutants. H3K36me3 levels were increased at E14.5Bip genes in Ezh2-cKO^{1,2-BP} mutant cells, as compared to wild-type cells
bodies, upregulation of mRNA levels in EedKO wild-type and EedKO Supplementary Note). We then analyzed the transcriptional upregulation of bipartite genes in full Ezh1−KO and Ezh2−KO cKO embryos, we used EedKO mouse ESCs in which the H3K27me3 mark is removed genome wide. We carried out RNAPII-S2P ChIP–seq and mRNA-seq in mouse ESCs in which the H3K27me3 mark is removed genome wide (Extended Data Fig. 7e). To overcome the unfeasibility of obtaining large amounts of cells from Ezh2-cKO embryos, we used EedKO mouse ESCs in which the H3K27me3 mark is removed genome wide. We carried out RNAPII-S2P ChIP–seq and mRNA-seq in wild-type and EedKO ESCs. For genes carrying H3K27me3 in gene bodies, upregulation of mRNA levels in EedKO cells correlated with a modest but significant increase of RNAPII-S2P signals in the TES region, compared with wild-type ESCs (Extended Data Fig. 7f,g and Supplementary Note). We then analyzed the transcriptional upregulation of bipartite genes in full Ezh1-KO;Ezh2-KO and Ezh2 catalytically inactive Ezh1-KO;Ezh2Δ722Δ722 mutant ESCs and found that the H3K27me3 mark itself on the gene body, rather than recruitment of Pc proteins, was required for the inhibition of bipartite gene productive transcription (Supplementary Note and Extended Data Fig. 7h). Taken together, these results indicate that the Pc-dependent H3K27me3 marking of the gene bodies of bipartite genes inhibits productive mRNA elongation.

To further support these findings, we selectively depleted the H3K27me3 mark from specific bipartite gene bodies and analyzed its acute effect on productive mRNA transcription. We developed an ex vivo short-term culture of E12.5 neurons from bulk hindbrain tissue; in this system, we observed no H3K27me3 depletion from bipartite gene bodies normally observed in long-term (1 week) hindbrain and cortical neuron embryonic cultures (Extended Data Figs. 2a and 9a). Overexpression of the catalytically ‘dead’ Cas9 (dCas9) fused to the H3K27me3 demethylase UTX (Kdm6a; dCas9–UTX) resulted in the selective decrease of H3K27me3 from the bipartite gene body (that is, Fos; Extended Data Fig. 9b–d). Quantification of mRNA levels confirmed that dCas9–UTX targeted to gene bodies of bipartite genes (Fos and Egr1) caused significant transcriptional upregulation of these genes (Fig. 5c), whereas dCas9–UTX targeted to non-bipartite gene bodies (Actb and Gapdh) did not affect gene expression (Extended Data Fig. 9e).

Together, these results indicate that the H3K27me3 histone mark on gene bodies of bipartite genes interferes with the production and accumulation of mature mRNA from the bipartite active promoters.

The bipartite signature regulates the rapidity and amplitude of transcriptional response to stimuli. Next, we asked whether the bipartite state might still allow rapid stimulus-dependent inducibility of IEGs, and whether bipartite or bivalent IEGs would display distinct transcriptional responses. We FACS-isolated cells from E14.5 hindbrain bulk tissue and treated them with 55 mM KCl for 8 or 30 min. KCl-mediated depolarization of cultured neurons results in an increase of intracellular calcium signaling and phosphorylation of CREB, a readout of stimulus-dependent transcription, at IEG promoters and is widely used to mimic the transcriptional response to a wide range of sensory stimuli. While an 8-min KCl treatment caused rapid induction of the bipartite Fos and Egr1 IEGs, the bivalent Junb IEG (Extended Data Fig. 10a) was not induced;
Fig. 5 | Polycomb marking of gene bodies inhibits productive mRNA processing and regulates rapidity and amplitude of transcriptional response to stimuli. a, Violin plots of gene body H3K27me3 and H3K27ac levels of bipartite genes (E14.5Bip; n=100; Methods) in control K20<sup>lox/lox</sup> (WT; green) and Ezh2-cKO<sup>lox/lox</sup> (purple) conditional knockout (cKO) E14.5 hindbrain cells. b, MA plot comparing productive Smart-seq2 mRNA levels of E14.5Bip genes (red dots; n=100; Methods) in control and cKO Ezh2<sup>−/−</sup>-cKO (purple) conditional knockout (cKO) E14.5 hindbrain cells. c, MA plot comparing productive Smart-seq2 mRNA levels of E14.5Bip genes (red dots; n=100; Methods) in control and cKO Ezh2<sup>−/−</sup>-cKO (purple) conditional knockout (cKO) E14.5 hindbrain cells. d, Violin plots, show log<sub>2</sub> fold changes of enhancer chromatin accessibilities in E14.5 Ezh2-cKO (Ezh2-cKO<sup>lox/lox</sup>) homozygous mutant neurons as compared to heterozygous control neurons; activity-dependent Fos-binding enhancers, normally open only at P4 (n=85), gain precocious accessibility when compared to all the remaining enhancers only open at P4 (n=3,882). e, Fos, Egr1, Arc and Klf4 mRNA levels were measured by quantitative PCR with reverse transcription (RT-qPCR) in short-term cultured E12.5 hindbrain (HB) neurons overexpressing control dCas9 or dCas9-UTX targeted to Fos (n=7 biologically independent neuron cultures) or Egr1 (n=6 biologically independent neuron cultures) gene bodies. gRNA, guide RNA. f, Fos and Egr1 mRNA levels by RT-qPCR in serum-starved WT and Eed<sup>−/−</sup> mouse ESCs treated with a low (1%) or high (10%) concentration of FCS for 8 min (n=4 biologically independent cultured cells). Treatment with 1% FCS was not sufficient to induce rapid transcriptional responses in WT, unlike in Eed<sup>−/−</sup>, ESCs; while 10% FCS induced rapid transcriptional responses in both WT and mutant, although levels were higher in Eed<sup>−/−</sup> ESCs. g, Summary of Pc-dependent regulation of rapidity and amplitude of bipartite gene transcriptional response to environmental stimuli with distinct strengths. In a and d, plots extend from the data minima to the maxima, the white dot indicates the median, the box shows the interquartile range and whiskers extend to the most extreme data point within 1.5 times the interquartile range. P values are from paired two-sided Wilcoxon’s tests. In e and f, the median expression is indicated by bars. P values are from Welch’s two-sample two-sided t-tests.
however, its transcripts could be detected after 30 min (Extended Data Fig. 10b). Thus, if developing neurons become exposed to a relevant signal, the bipartite signature at the Fos and Egr1 loci might still allow for rapid inducibility, whereas the bivalent state constrains the Junb IEG to a slower response and only in the presence of prolonged stimulation.

We next evaluated the amplitude of the transcriptional response of bipartite IEGs to distinct strengths of the same signal. We used serum treatment after starvation in mouse ESCs, a well-known model to rapidly induce expression of IEGs. Fos and Egr1 carried the bipartite signature also in mouse ESCs (Extended Data Fig. 3d). We treated serum-starved wild-type and EedKO ESCs with a low (1%) or high (10%) concentration of fetal calf serum (FCS) for a short (8 min) or a longer (16 min) time of exposure and quantified Fos and Egr1 transcriptional induction (Fig. 5f and Extended Data Fig. 10c). Treatment with 10% FCS could induce a rapid (within 8 min) Fos and Egr1 transcriptional response in both wild-type and EedKO backgrounds; however, the amplitude of the Fos and Egr1 transcriptional responses was higher in EedKO than wild-type ESCs (Fig. 5f). Furthermore, lowering the concentration of the stimulus by tenfold (that is, treating with 1% FCS) was not sufficient to elicit a transcriptional response after an 8-min treatment in wild-type ESCs but caused significant Fos and Egr1 induction in the EedKO background (Fig. 5f). In wild-type ESCs, the bipartite Fos and Egr1 IEGs could only be induced after prolonged exposure (that is, 16 min) to 1% FCS (Extended Data Fig. 10c).
In summary, H3K27me3 marking of bipartite IEG gene bodies, while still allowing for rapid induction, regulates the amplitude of the transcriptional response to relevant stimuli. Moreover, Pc marking of gene bodies of bipartite stimulus-response genes may establish a transcriptional threshold to prevent rapid productive induction of IEGs in response to suboptimal and/or nonphysiologically relevant levels of environmental stimuli (Fig. 5g).

Mechanism of stimulus-dependent transition of bipartite to active chromatin. Negative elongation factor (NELF) negatively regulates transcriptional elongation by pausing RNAPII at TSSs. Stimulus-dependent NELF removal from IEG promoters causes release of paused RNAPII into elongation. We found that H3K27me3 on gene bodies inhibits transcriptional elongation in bipartite genes in part by interfering with stimulus-dependent NELF release (Fig. 6a and Supplementary Note). Moreover, Ezh1/ Ezh2 removal caused a reduction in levels of Ring1b in the gene body of bipartite genes (Fig. 6b), correlating with a significant increase of bipartite gene body, although not promoter, accessibility in Ezh1-KO/Ezh2-KO mouse ESCs (Fig. 6c and Supplementary Note). Such decompaction of bipartite gene bodies was not only merely correlated with increased transcription, but was at least partially caused by the removal of H3K27me3 (Fig. 6d,e and Supplementary Note).

As for the transition from a bipartite to an active state, we reasoned that stimulus-dependent posttranslational modification of transcription factors prebound to promoters could be involved, and in turn induce an increase of H3K27ac, decrease of H3K27me3 and gain of productive transcription (Extended Data Fig. 6a). CREB phosphorylation is rapidly increased in response to neuronal activity and/or other environmental stimuli and induces CREB-binding protein-dependent H3K27ac increase and transcription of IEGs. Indeed, phosphoCREB (pCREB) levels increased in the promoter regions of genes that were bipartite at E14.5 and became active at P4 (Fig. 7a; Bip → Exp), including neuronal activity-induced IEGs such as Fos and Egr1 (Fig. 3e and Extended Data Fig. 6b). This correlated with the resolution of the bipartite signature and productive transcription (Fig. 7a and Extended Data Fig. 6a,d,e).

Are strong inducing stimuli (for example, neuronal activity) able to resolve the bipartite epigenetic state? By treating E12.5 short-term cultured hindbrain neurons with 55 mM KCl, after overnight incubation with a cocktail of neuronal activity blockers (TGN cocktail = tetrodotoxin + D-AP5 + NBQX; Methods), the H3K27me3 mark was removed from IEG gene bodies (Fig. 7b). Notably, the decrease of the H3K27me3 mark was detectable as early as 8 min after KCl treatment (Fig. 7b), showing that H3K27me3 removal starts very rapidly after exposure to the inducing stimuli. In addition, treatment of embryonic neurons with a TGN cocktail prevented the removal of H3K27me3 from IEG gene bodies in long-term hindbrain neuron culture (Fig. 7c; also see above and Extended Data Fig. 9a). This indicates that the removal of H3K27me3 from IEG gene bodies is rapid and stimulus dependent.

Furthermore, treatment with GSK-J4, an inhibitor of H3K27me3 demethylases (that is, UTx (Kdm6a), Jmd3 (Kdm6b)) prevented neuronal activity-dependent gene body H3K27me3 removal (Fig. 7d). Similarly, inactivation of Jmd3 inhibited, at least partially, gene body H3K27me3 removal from the E14.5Bip genes that became active at perinatal/postnatal stages (Fig. 7e, Supplementary Table 1 and Supplementary Note). These results indicate that the stimulus-dependent removal of H3K27me3 from IEG gene bodies requires active demethylation. In addition, GSK-J4 treatment prevented the rapid transcriptional induction of bipartite IEGs after short (8 min) exposure to the inducing stimulus (Fig. 7f). Taken together with our previous observation that, in the absence of the H3K27me3 mark in Ecd ESCs, the amplitude of the rapid bipartite IEG transcriptional response upon short exposure (8 min) to inducing stimuli (that is, FCS) is enhanced as compared to that in wild-type control (Fig. 5i), these results indicate that stimulus-dependent gene body H3K27me3 mark removal is essential to achieve rapid and sizeable transcriptional induction of bipartite IEGs.

On the other hand, after prolonged exposure (that is, 60 min) to the inducing stimulus, GSK-J4-treated neurons showed transcriptional upregulation of bipartite IEGs, even though mRNA levels remained significantly lower than those in control neurons (Fig. 7f). Thus, in the event of incomplete H3K27me3 mark removal from the gene body, while rapid bipartite IEG mRNA induction is impaired, transcripts can nonetheless accumulate over time upon prolonged stimulation, albeit never reaching optimal levels.

We then tested the requirement of de novo promoter H3K27 acetylation in activity-dependent removal of the gene body H3K27me3. We treated E12.5 short-term cultured neurons with KCl in the presence of A-485, an inhibitor of H3K27 acetyltransferase p300/CREB-binding protein. A-485 inhibited KCl-dependent increase of promoter H3K27ac levels and prevented the removal of the

---

**Fig. 7 | Mechanism of stimulus-dependent transition from bipartite to active chromatin.** a Violin plots visualizing log2 fold changes of promoter pCREB levels of E14.5Bip genes from E14.5 to P4; E14.5Bip genes that became expressed (n = 20 genes), bivalent (n = 25 genes) or remained bipartite (n = 55 genes) at P4 were compared. Plots extend from the data minima to the maxima, the white dot indicates the median, the box shows the interquartile range and whiskers extend to the most extreme data point within 1.5 times the interquartile range. b After overnight (o/n) treatment with TGN cocktail (inhibitors of sodium channel, NMDA and AMPA receptors), E12.5 short-term cultured hindbrain neurons were treated with 55 mM KCl for short (8 min) and prolonged (60 min) time courses at day 2. H3K27me3 decrease was detectable as early as 8 min after KCl treatment. b-d,f-h, qPCR quantification of mRNA, H3K27ac and H3K27me3 levels from n = 3 biologically independent neuron cultures (bars indicate the median). c, Shorter (1d) or longer (4d) E12.5 hindbrain neuron cultures in the presence or absence of the TGN cocktail. d, E12.5 short-term cultured hindbrain neurons treated by KCl for 1 h in the absence or presence of GSK-J4 at day 2. KCI treatment caused H3K27me3 removal from bipartite IEGs through active demethylation. e, MA plot comparing gene body H3K27me3 levels of E14.5Bip genes that became expressed at P4 (n = 20 genes; a) between E18.5 Jmd3-KO and WT hindbrain cells. f, Short (8 min) or prolonged (60 min) KCl treatment of E12.5 short-term cultured hindbrain neurons in the absence or presence of GSK-J4 at day 2, showing that inhibition of H3K27me3 removal prevented rapid induction of bipartite IEGs. g, KCl treatment (1h) of E12.5 short-term cultured hindbrain neurons in the absence or presence of A-485 at day 2. h, Short (8 min) KCl treatment of E12.5 short-term cultured hindbrain neurons in the absence or presence of A-485 at day 2. Inhibition of de novo promoter H3K27ac prevented rapid KCl-dependent induction of bipartite IEGs (n = 3; P values are from a two-sided t-test). i, Diagram of H3K27 demethylation (GSK-J4) or acetyltransferase (A-485) inhibitor treatments on neuronal activity-dependent transcriptional induction of bipartite IEGs after short (8 min) or prolonged (60 min) exposure to strong inducing stimulus (55 mM KCl). j, KCl treatment (1h) of E12.5 short-term cultured hindbrain neurons in the absence or presence of flavopiridol (Flav) at day 2. KCl-induced H3K27me3 removal of bipartite IEGs appears to be dependent on de novo promoter H3K27ac but not on transcriptional elongation. k, Scatter plot showing log2 fold changes of gene body H3K27ac (x axis) and H3K27me3 (y axis) levels upon overnight TSA treatment in short-term cultured E12.5 hindbrain neurons. E14.5Bip gene distribution is mapped. Colors indicate the log fold changes of mRNA levels of E14.5Bip genes. The genome browser view of Fos in TSA-treated short-term cultured hindbrain neurons shows H3K27me3 levels were reduced by expansion of H3K27ac in the coding region (green highlight). In b–d, g and j, P values are from analysis of variance followed by Tukey’s HSD posthoc tests. In f and h, P values are from Welch’s two-sample two-sided t-tests.
H3K27me3 mark from bipartite IEG gene bodies (Fig. 7g), indicating that gene body H3K27me3 removal requires stimulus-dependent de novo promoter H3K27 acetylation. Furthermore, A-485 treatment prevented rapid induction of bipartite IEGs after short-time (that is, 8 min) exposure to KCl (Fig. 7h), similarly to GSK-J4 treatment (Fig. 7f), indicating that fast bipartite IEG transcriptional induction requires de novo H3K27 acetylation and rapid removal of the gene body H3K27me3 mark through active demethylation (Fig. 7i and Supplementary Note). Moreover, the KCl-dependent gene body H3K27me3 removal is not merely the consequence
Fig. 8 Polycomb-dependent regulation of stimulus-response genes during development. During development, a subset of stimulus-response genes displays a bipartite chromatin signature that carries active H3K27ac/H3K4me2 on promoters but repressive H3K27me3 on gene bodies. The bipartite signature originates from H3K27me3/H3K4me2+ bipartite chromatin maintaining a transcriptionally poised state and nonproductive transcription. Distinct transcription factors cause partial resolution of a subset of bivalent poised promoters of stimulus-response genes into the bipartite state. Active promoters of bipartite genes carry actively initiating RNAPII (high 8WG16, SSP and S7P); however, productive mRNA processing and elongation in gene bodies are maintained at a low rate (low RNAPII-S2P and H3K36me3) due to inhibition of stimulus-dependent NELF release and chromatin compaction by Pc (Ezh2/2, Eed)-dependent H3K27me3 (inhibition sign). Pc-dependent marking on gene bodies also inhibits spreading of H3K27ac and accessibility in bipartite gene body regions (mutual inhibition signs). The bipartite state is dynamic and could revert to bivalency. The bipartite signature maintains the potential for fast induction by relevant/strong stimuli while preventing inappropriate induction by nonrelevant/weak stimuli. Inducing stimuli cause de novo promoter H3K27 acetylation, which causes H3K27 demethylase (Kdm6; that is, UTX and Jmjd3)-dependent rapid removal of the Pc mark from bipartite gene bodies (inhibition sign) and fast transcriptional response by transcription factors.

of transcriptional elongation but it is at least partly driven by the de novo promoter acetylation per se (Fig. 7) and Supplementary Note.

Lastly, treatment of E12.5 short-term cultured neurons with the histone deacetylase inhibitor trichostatin A (TSA) resulted in the spreading of H3K27ac into the bipartite gene bodies and H3K27me3 removal, increase of mRNA levels and resolution of the bipartite signature into an active state (Fig. 7k). Together with the analysis of E14.5 Ezh2-CKO hindbrain cells (Fig. 5a), and the finding that bipartite genes can revert to bivalency during development (Fig. 3d and Extended Data Fig. 6a), we propose that a dynamic reciprocal balance between the H3K27ac and H3K27me3 marks maintains the bipartite signature.

In summary, stimulus-dependent increase of promoter H3K27ac causes active and rapid H3K27me3 removal from the gene body and release of the elongation barrier, switching from the bipartite to the productive active transcription state.

Discussion

During development, the response to environmental signals requires rapid, stimulus-dependent, transcriptional responses through the induction of IEGs, whose gene products in turn regulate the activation of specific LRGs, driving cell-type-specific differentiation schedules. How chromatin states and epigenetic regulation contribute to the timely and rapid activation of stimulus-induced developmental transcriptional programs is poorly understood. Here, we discovered an unusual Pc-dependent bipartite chromatin signature at stimulus-response IEGs before their transcriptional induction in developing neurons, whereas LRGs were preferentially maintained in a bivalent chromatin state. Moreover, we found that the bipartite state is not an exclusive feature of developing neurons, but it is generally present in developing cell types and in ESCs. The bipartite state originates from the bivalent state and is dynamic during development, reverting to bivalency or resolving into rapid activation (Fig. 8). Bipartite genes carry an active promoter and the Pc-dependent H3K27me3 mark on the gene body, which inhibits RNAPII transcriptional elongation regulating the transition into stimulus-dependent productive transcription of bipartite genes (Fig. 8). We demonstrate that this unique chromatin signature provides a suitable epigenetic structure to modulate the rapidity and amplitude of the transcriptional response of inducible IEGs to distinct stimuli during development while inhibiting IEG productive transcription in response to suboptimal and/or nonphysiologically significant levels of environmental stimuli (Fig. 5g). Additional discussion can be found in the Supplementary Discussion.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-021-00789-z.

Received: 5 October 2020; Accepted: 19 January 2021; Published online: 18 February 2021

References

1. Fowler, T., Sen, R. & Roy, A. L. Regulation of primary response genes. Mol. Cell 44, 348–360 (2011).
2. West, A. E. & Greenberg, M. E. Neuronal activity-regulated gene transcription in synapse development and cognitive function. Cold Spring Harb. Perspect. Biol. https://doi.org/10.1101/cshperspect.a005744 (2011).
3. Greenberg, M. E. & Ziff, E. B. Stimulation of T3 cells induces transcription of the c-fos proto-oncogene. Nature 311, 433–438 (1984).
4. Malik, A. N. et al. Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. Nat. Neurosci. 17, 1330–1339 (2014).
5. Vierbuchen, T. et al. AP-1 transcription factors and the RAF complex mediate signal-dependent enhancer selection. Mol. Cell 68, 1067–1082 (2017).
6. Stroud, H. et al. An activity-mediated transition in transcription in early postnatal neurons. Neuron 107, 874–890 (2020).
7. Mayer, A., Landry, H. M. & Churchman, L. S. Pause & go: from the discovery of RNA polymerase pausing to its functional implications. Curr. Opin. Cell Biol. 66, 72–80 (2017).
8. Yap, E. L. & Greenberg, M. E. Activity-regulated transcription: bridging the gap between neural activity and behavior. Neuropharmacology 100, 330–348 (2018).
9. Lonze, B. E. & Ginty, D. D. Function and regulation of CREB family transcription factors in the nervous system. Neuropharmacology 35, 505–623 (2002).
10. Toth, A. B., Shum, A. K. & Prakriya, M. Regulation of neurogenesis by calcium signaling. Cell Calcium 59, 124–134 (2016).
11. Ginty, D. D., Glowacka, D., Bader, D. S., Hidaka, H. & Wagner, J. A. Induction of immediate early genes by Ca+1 influx requires CAMP-dependent protein kinase in PC12 cells. J. Biol. Chem. 266, 17454–17458 (1991).
12. Greenberg, M. E., Greene, L. A. & Ziff, E. B. Nerve growth factor and epidermal growth factor induce rapid transient changes in proto-oncogene transcription in PC12 cells. J. Biol. Chem. 260, 14101–14110 (1995).
13. Erzurumlu, R. S., Murakami, Y. & Riiju, F. M. Mapping the face in the somatosensory brainstem. Nat. Rev. Neurosci. 11, 252–263 (2010).

90

390
14. Kitazawa, T. & Rijli, F. M. Barrelette map formation in the prenatal mouse brainstem. *Curr. Opin. Neurobiol.* 53, 210–219 (2018).

15. Erzurumlu, R. S. & Gaspar, P. Development and critical period plasticity of the barrel cortex. *Eur. J. Neurosci.* 35, 1540–1553 (2012).

16. Hrvatin, S. et al. Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci.* 21, 120–129 (2018).

17. Tyssowski, K. M. et al. Different neuronal activity patterns induce different gene expression programs. *Neuron* 98, 530–546 (2018).

18. Valles, A. et al. Genomewide analysis of rat barrel cortex reveals time- and layer-specific mRNA expression changes related to experience-dependent plasticity. *J. Neurosci.* 31, 6140–6158 (2011).

19. Mohn, F. et al. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* 30, 755–766 (2008).

20. Ferrai, C. et al. RNA polymerase II primes Polycomb-repressed developmental genes throughout terminal neuronal differentiation. *Mol. Syst. Biol.* 13, 946 (2017).

21. Hirabayashi, Y. et al. Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. *Neuron* 63, 600–613 (2009).

22. Aranda, S., Mas, G. & Di Croce, L. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* 1, e1500737 (2015).

23. Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome regulation by Polycomb and Trithorax: 70 years and counting. *Cell* 171, 34–57 (2017).

24. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326 (2006).

25. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science* https://doi.org/10.1126/science.aal2913 (2017).

26. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* 352, aa9780 (2016).

27. Bonnefont, J. et al. Cortical neurogenesis requires Bcl6-mediated transcriptional repression of multiple self-renewal-promoting extrinsic pathways. *Neuron* 103, 1096–1108 (2019).

28. Chen, F. X., Smith, E. R. & Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 19, 464–478 (2018).

29. Brookes, E. & Pomar, A. Modifications of RNA polymerase II are pivotal in regulating gene expression states. *EMBO Rep.* 10, 1213–1219 (2009).

30. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.* 23, 771–777 (2016).

31. Brookes, E. et al. Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. *Cell Stem Cell* 10, 157–170 (2012).

32. Stock, J. K. et al. Ring1-mediated ubiquitination of H2A restrains RNA polymerase II at bivalent genes in mouse ES cells. *Nat. Cell Biol.* 9, 1428–1435 (2007).

33. Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* 10, 697–708 (2009).

34. Blackledge, N. P., Rose, N. R. & Klose, R. J. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. *Nat. Rev. Mol. Cell Biol.* 16, 643–649 (2015).

35. Shen, X. et al. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol. Cell* 32, 491–502 (2008).

36. Schoeftner, S. et al. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* 25, 3110–3122 (2006).

37. Lavarone, E., Barbieri, C. M. & Pasini, D. Dissecting the role of H3K27 acetylation and methylation in PRC2-mediated control of cellular identity. *Nat. Commun.* 10, 1679 (2019).

38. Kim, T. K. et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187 (2010).

39. Schoekowitsch, K. et al. Enhancer RNA facilitates NELF release from immediate early genes. *Mol. Cell* 56, 29–42 (2014).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021
**Methods**

**Mating scheme.** To obtain E10.5 and E14.5 Krox20::Cre;R26loxP/loxP (K20loxP) embryos, the Krox20::Cre transgenic mouse line was crossed with the R26loxP reporter mouse line (The Jackson Laboratory, 007906). To obtain E14.5, E18.5 and P4 Drg11::Cre;R26loxP/loxP (mCherry) (Drg11Cre/+(mCherry)) mice, Drg11::Cre transgenic mouse line, the R26loxP reporter mouse line and the R26loxP (transgenic mouse line, Supplementary Methods) were crossed (Extended Data Fig. 1d). To obtain E18.5 Drg11::Cre;R26loxP/loxP;E2GFP (Drg11Cre/+(loxP-E2GFP)) transgenic mouse line, the Drg11::Cre transgenic mouse line, the R26loxP E2Cre transgenic mouse line, the R26loxP reporter mouse line and the R26loxP (transgenic mouse line, Supplementary Methods) were crossed (Extended Data Fig. 1d). To obtain E18.5 Drg11::Cre;R26loxP/loxP;E2GFP (Drg11Cre/+(loxP-E2GFP)) transgenic mouse line, the Drg11::Cre transgenic mouse line, the R26loxP E2Cre transgenic mouse line, the R26loxP reporter mouse line and the R26loxP (transgenic mouse line, Supplementary Methods) were crossed (Extended Data Fig. 1d). To obtain E18.5 Drg11::Cre;R26loxP/loxP;E2GFP (Drg11Cre/+(loxP-E2GFP)) transgenic mouse line, the Drg11::Cre transgenic mouse line, the R26loxP E2Cre transgenic mouse line, the R26loxP reporter mouse line and the R26loxP (transgenic mouse line, Supplementary Methods) were crossed (Extended Data Fig. 1d). To obtain E18.5 Drg11::Cre;R26loxP/loxP;E2GFP (Drg11Cre/+(loxP-E2GFP)) transgenic mouse line, the Drg11::Cre transgenic mouse line, the R26loxP E2Cre transgenic mouse line, the R26loxP reporter mouse line and the R26loxP (transgenic mouse line, Supplementary Methods) were crossed (Extended Data Fig. 1d).

**Processing of these cells was adapted for further analyses (RNA extraction and ChIP–seq of H3K27me3; see below and Supplementary Methods).**

**KCl, trichostatin A, TDN cocktail, GSK-J4, A-485 and flavopiridol treatment.** Ex vivo cultured E12.5 hindbrain neurons were treated with 1 μM TSA (MBI, JM-1606-1) at the culture day 1 and incubated for 16 h. For KCl treatment, cultured hindbrain neurons were treated with a cocktail of neuronal activity blockers (TDN cocktail = 1 μM tetrodotoxin (TOCRIS, 1069) + 100 μM D-AP5 (Sigma, A8054) + 20 μM NBQX (TOCRIS, 0373) at day 2 for an overnight incubation, and 55 mM KCl-containing medium was treated at day 2 in the presence or absence of 35 μM GSK-J4 (Sigma, SML0701), 50 μM A-485 (TOCRIS, 6387) or 10 μM flavopiridol (Sigma, F3055) after a rinse. For Drg11Cre/+(loxP-E2GFP) cultivated hindbrain neurons, tdTomato+ neurons were sorted immediately after the KCl treatment. Processing of these cells was adapted for further analyses (mRNA-seq, ATAC-seq and ChIP-seq; see below and Supplementary Methods).

**Serum shock of mouse ESCs.** Wild-type and E2GFP mouse ESCs were cultured up to 80% confluence in normal culture medium (Supplementary Methods) and were subsequently serum starved overnight in the culture medium that did not contain FCS. Serum-starved ESCs were treated by a low (1%) or high (10%) concentration of FCS for a short (6 h) or a longer (12 h) time. After reaction, total RNA was immediately extracted by RNeasy Mini Kit (QIAGEN, 79254) according to the manufacturer’s protocol, and RT–qPCR was conducted (Supplementary Methods).

**Sample preparation and RNA isolation and sequencing.** For RNA-seq experiments, total RNA was extracted by a Single Cell RNA Purification Kit (NORGAN, 51800) with genomic DNA digestion using an RNase-Free DNase I Set (QIAGEN, 25710) according to the manufacturers’ protocols. Library preparation protocols for poly A* mRNA (Smart-seq2 protocol) and total RNA (Ovation SoLo RNA-seq System), as well as for single-cell RNA-seq (10x Genomics), are described in Supplementary Methods.

**Sample preparation and chromatin immunoprecipitation and sequencing.** Cells were cross-linked with 1% formaldehyde for 10 min at 25°C and quenched with 125 mM glycine for 5 min at room temperature. To achieve the sequencing of chromatin immunoprecipitated from small amounts of cells, preparation of the ChIP-seq library was mostly done by ChiPmentation protocol. Cells were lysed in sonication buffer (10 mM Tris-HCl (pH 8), 5 mM EDTA, 0.5% SDS), 1x proteinase inhibitor cocktail (complete EDTA-free; Roche, 0469312001) on ice, and sonicated using the Covaris machine to obtain DNA fragments between 150 bp and 500 bp. The supernatant was transferred to a new tube, diluted with equilibration buffer (10 mM Tris-HCl (pH 8), 1 mM EDTA, 140 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate and 1× proteinase inhibitor cocktail). Chromatin solutions were incubated overnight at 4°C with antibodies. The next day, protein G magnetic beads (Dynabeads Protein G; Thermo Fischer, 10004D) were washed and the incubation was continued for 2 h at 4°C. The beads were then washed and resuspended in tagmentation buffer (10 mM Tris-HCl (pH 8) and 5 mM MgCl2) containing Tagment DNA Enzyme from the Nextera DNA Sample Prep Kit (Illumina, FC-121-1030) and incubated at 37°C for 10 min. The beads were washed, and DNA was eluted from the beads with elution buffer (10 mM Tris-HCl (pH 8), 5 mM EDTA, 300 mM NaCl, 0.5% SDS and protease K) at 65°C. DNA was purified with Spri AMPure XP beads (Beckman Coulter; sample to beads ratio of 1:2) and eluted in 10 mM Tris-HCl (pH 8). Libraries were prepared in a 50-μl reaction (1X KAPA HiFi Hot Start Ready Mix and 0.8 μM primers). Enriched libraries were purified with size selection using SPRI Ampure XP beads (sample to beads ratio of 1:6.6) to remove long fragments and recovering the remaining DNA (sample to beads ratio of 1:2). Sequencing was performed on an Illumina HiSeq 2500 (50 bp read length, single-end). The ChIP-seq protocol was optimized for different experiments (for example, FACS-sorted cells, bulk tissue, cultured cells, prefixed tissue and sequential ChIP-seq; Supplementary Methods).

**Sample preparation and assay for transposase-accessible chromatin.** ATAC-seq experiments were performed as described previously with minor modifications. For each experiment, 50,000–70,000 cells were used. Two independent biological replicates were prepared. For the detailed protocol of ATAC-seq, see Supplementary Methods.

**Reference genome and annotation.** The mouse GRCm38/mm10 genome assembly was used as reference. The most variable TSS in promoter chromatin assembly was used as reference. The most variable TSS in promoter chromatin assembly was used as reference.
Articles

Nature Genetics

were performed with ‘edgeR’ using ‘glmQLFit’. Non-barrelette ARGs (nbARGs) were specific to barrelette neurons (bsARGs) were Activity-response genes. In Supplementary Methods. For the initial 100 top-scoring genes with t-SNE (E10.5 t-SNE) and the top 300 genes for the combined t-SNE map (E14.5/E18.5/ P4 combined t-SNE), 2D densities for bipartite and bivalent genes were estimated with ‘ide2’ from MASS and visualized as contour lines (Extended Data Fig. 5g,l). We calculated Euclidean distances of genes between P4 and E14.5 on the original eight-dimensional space consisting of normalized log (RPKM) counts of ATAC, H3K27me3, H3K27ac and H3K4me2 on the P and GR regions and colored the E14.5 t-SNE by this distance (Extended Data Fig. 5h). The 100 top-scoring genes for bipartiteness at E14.5 were divided into three groups: genes that became expressed at P4 (RPKM ≥ 2; 20 genes), genes that became bivalent (move into the bivalent contour at P4; Fig. 4d; 25 genes) and those that remain bipartite (55 genes; Extended Data Fig. 6a).

Gene sets for comparison to E14.5 bipartite genes. For Fig. 4, we first selected the 100 top-scoring bipartite and bivalent genes at E14.5 (E14.5Bip and E14.5Biv) and excluded three genes contained in both sets. Control sets of the same number of genes (97) were then created using ‘swissknife’. E14.5AcP genes were sampled from all genes except the top 400 E14.5Bip genes and E14.5Biv genes, to have similar H3K27ac distribution in P as E14.5Bip. E14.5mRNA-Low genes that matched E14.5Bip in log, RPKM mRNA expression were sampled similarly from all genes excluding E14.5Bip, the top 400 E14.5Bip genes and E14.5AcP. Finally, two sets were sampled from the bottom and top 30% of genes ordered by mRNA expression, excluding any of the genes already contained in the previous sets. For the E10.5 samples, sets with a total of 99 genes were similarly created, excluding 1 gene that was common between the top 100 bipartite and bivalent genes (Extended Data Fig. 4e).

For Extended Data Fig. 7a,b, Entrez identifiers from the top 100 E14.5Bip genes were mapped to Ensemble IDs using biomart®, resulting in 90 successfully mapped identifiers. A control set of the same size with matching spliced transcript abundance, excluding the top 400 E14.5Bip genes, was then randomly sampled.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All raw sequencing data and processed data used for this study are available through ArrayExpress and will be released to the public without restrictions: mRNA-seq (Smart-seq2), E-MTAB-8314; total RNA-seq (Solo RNA-seq), E-MTAB-8311; ChIP-seq (ChIPmentation), E-MTAB-8317; ATAC–seq, E-MTAB-8313; 4C–seq, E-MTAB-8295; and single-cell RNA-seq (10x Genomics), E-MTAB-8296. All raw sequencing data and processed data used for this study are available through ArrayExpress and will be released to the public without restrictions: mRNA-seq (Smart-seq2), E-MTAB-8314; total RNA-seq (Solo RNA-seq), E-MTAB-8311; ChIP-seq (ChIPmentation), E-MTAB-8317; ATAC–seq, E-MTAB-8313; 4C–seq, E-MTAB-8295; and single-cell RNA-seq (10x Genomics), E-MTAB-8296. FACs gating strategies/source data are presented in Supplementary Figs. 1–4. Public sequencing datasets were obtained from the Gene Expression Omnibus and ENCODE as follows: mouse cortical culture (GSE21161 and GSE60192), mouse embryonic forebrain (GSE93011 and GSE52386), mouse adult cortical excitatory neuron (GSE6317), mouse ESCs (GSE36114 and GSE94250), mouse ESCs for Eh2KO experiments (GSE16603), mouse E14.5 liver tissues (GSE74412, GSE2407, GSE2615 and GSE28262), ENSCR002HKE, and E10.5 mouse NCCs isolated from the frontal nasal process (GSE94937).

Code availability

Computational analyses were performed in R using the mentioned publicly available packages (Methods, Reporting Summary and Supplementary Methods). The custom tool monaLisa (v0.12.8), used for motif enrichment, can be found on GitHub at https://github.com/fmcompo/mona-lisa/. The custom tool swissknife (v0.10.0) is available on https://github.com/fmcompo/swissknife/.

References

40. Voiculescu, O. et al. Hndbrain patterning: Krox20 couples segmentation and specification of regional identity. Development 128, 4967–4978 (2001).
41. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).
42. Bechara, A. et al. *Hoxa2* selects barrelette neuron identity and connectivity in the mouse somatosensory brainstem. *Cell Rep.* **13**, 783–797 (2015).

43. Moreno-Juan, V. et al. Prenatal thalamic waves regulate cortical area size prior to sensory processing. *Nat. Commun.* **8**, 14172 (2017).

44. Luche, H., Weber, O., Nageswara Rao, T., Blum, C. & Fehling, H. J. Faithful activation of an extra-bright red fluorescent protein in knock-in Cre-reporter mice ideally suited for lineage tracing studies. *Eur. J. Immunol.* **37**, 43–53 (2007).

45. Di Meglio, T. et al. *Ezh2* orchestrates topographic migration and connectivity of mouse precerebellar neurons. *Science* **339**, 204–207 (2013).

46. Maheshwari, U. et al. Postmitotic Hoxa5 expression specifies pontine neuron positional identity and input connectivity of cortical afferent subsets. *Cell Rep.* **31**, 107767 (2020).

47. Seidler, B. et al. A Cre-loxP-based mouse model for conditional somatic gene expression and knockdown in vivo by using avian retroviral vectors. *Proc. Natl Acad. Sci. USA* **105**, 10137–10142 (2008).

48. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Methods* **10**, 1404–1408 (2013).

49. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

50. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

51. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Mol. Cell* **43**, e47 (2015).

52. O’Connell, J. & Hojsgaard, S. Hidden semi-Markov models for multiple observation sequences: the mhsmm package for R. *J. Stat. Softw.* **39**, 1 (2011).

53. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).

54. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Nat. Cell Biol.* **38**, 576–589 (2010).

55. Maaten, L. J. P. V. D. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

56. Venables, W. N. & Ripley, B. D. Modern Applied Statistics with S 4th edn (Springer, 2002).

57. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

Acknowledgements

F.M.R. wishes to dedicate this paper to the memory of P. Sassone-Corsi (1956–2020), dear friend and eminent scientist who made seminal contributions to immediate early gene and activator protein-1 transcriptional regulation and with whom F.M.R. discussed this project at an early stage. We thank A. Pombo for very useful discussion. We thank N. Vilain, S. Smallwood, D. Gaidatzis and the members of the Rijli group and FMI facilities for excellent technical support and discussion. We thank S. H. Orkin (Harvard Medical School) and A. Wutz (ETH Zurich) for the kind gifts of the *EZH2* mouse line and *Eed*−/− ESCs, respectively. The LAL-R26TVA-LacZ−/− line was a kind gift from D. Saur (Technische Universität Munchen). T.K. was supported by a Japan Society for the Promotion of Science fellowship, and O.I. was supported by an EMBO Long-Term fellowship. F.M.R. was supported by the Swiss National Science Foundation (31003A_149573 and 31003A_175776). This project has also received funding from the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant no. 810111-EpiCrest2Reg). F.M.R. and M.B.S. were also supported by the Novartis Research Foundation.

Author contributions

T.K. and F.M.R. conceived the study, designed experiments and analyzed experimental data. T.K. performed most of the experiments. H.K. carried out cell sorting. D.M. and T.K. and M.B.S. performed computational analysis. O.I. performed 4C–seq. S.K. carried out some ChIP–seq assays. S.D., H.G. and G.L.-B. contributed to Kir-OE mouse generation and characterization. N.M. analyzed the phenotype of Kir-OE mice. C.S. performed analysis of scRNA-seq. C.S. and P.P. contributed to the Solo RNA-seq analysis; T.K., D.M., and M.B.S. wrote the first draft. F.M.R. revised and wrote the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-021-00789-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-021-00789-z.

Correspondence and requests for materials should be addressed to F.M.R.

Peer review information *Nature Genetics* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Genetic strategy of barrelette neuron isolation and identification of activity response genes. a and b, Representative FACS gating for barrelette neurons (Supplementary Figs. 1–4). c and d, Intersectional strategies to FACS isolate E14.5, E18.5 or P4 postmitotic barrelette neurons (Drg11vPrV-ZsGreen+ (c), Drg11vPrV-tdTomato+ (d)) from ventral principal trigeminal nucleus (vPrV). (Supplementary Table 1, Supplementary Note). e, Intersectional strategy to FACS isolate Kir2.1(Kir)-mCherry overexpressing, neuronal activity-deprived, vPrV barrelette neurons (Drg11vPrV-Kir+; Supplementary Table 1, Supplementary Note).f, Volumes of PrV in control (K20tdTomato+/r2EGFP+) and Kir overexpressing (K20Kir+/r2EGFP+) mice (n = 3, biologically independent animals). g and h, P8 cytochrome oxidase (CO) staining in wild-type (WT; g) and K20Kir+ (h) mice. Representative images of n = 3 biologically independent animals. Scale bars: 200 μm. i–p, P10 barrelette neuron dendrite orientation by GFP expression after pseudotyped rabies virus (EnvA-SADΔG-GFP) injection in P3 thalamus in control Krox20::Cre;LslR26TVA-LacZ (K20TVA+/+) (j) and Kir overexpressing Krox20::Cre;R26Kir-mCherry;LslR26TVA-LacZ (K20TVA/Kir) (k) mice. Scale bars: 10 μm. Symmetry index (l) and surface ratio (p) are compared (Supplementary Methods). n = 8 (K20TVA+/+) and n = 6 (K20TVA/Kir) biologically independent animals were used, and j, k, n, o, are representative images. q and r, MA-plots comparing E18.5 and E14.5 mRNA levels in control Drg11vPrV-ZsGreen+ barrelette neurons (q), and E18.5 Drg11vPrV-Kir+ (Kir-OE) and E18.5 Drg11vPrV-ZsGreen+ wild-type (WT) barrelette neurons (r). 702 genes (green dots, q) increase their expression at E18.5 as compared to E14.5 (log2(fold change) > 1.5), while 102 genes (red dots, r) decrease their expression in E18.5 Kir-OE barrelette neurons (log2(fold change) < -1) (Methods). s, Identification of 56 bsARGs, t, Fractions of the 56 bsARGs (left) and 83 nbARGs (right) with H3K27me3- and H3K27me3+ profiles in E14.5 Drg11vPrV-ZsGreen+ barrelette neurons (see Fig. 1c). u, Scatterplots showing ATAC-seq (x axis) and H3K4me2 (y axis) signals on promoters (1 kb around TSS) in E14.5 Drg11vPrV-ZsGreen+ barrelette neurons. Dashed lines indicate thresholds corresponding to a 5% false discovery rate (FDR) (Methods) (see Fig. 1c). f, i, p, Bars indicate median and P values are from Welch’s two-sample two-sided t-tests. NS: not significant (P > 0.05).
Extended Data Fig. 2 | Chromatin profiles of IEGs carrying the bipartite signature. a, Fos, Egr1, Fosb and Nr4a3 genome browser views. ATAC (violet), H3K4me2 (yellow, E14.5 barrelette neurons), H3K4me3 (yellow, adult cortical neurons and cultured embryonic neurons), H3K27ac (red) and H3K27me3 (blue) are shown in E14.5 Drg11vPrV-ZsGreen/+ barrelette neurons, E15.5 embryonic cortical neural progenitors (NPC) and postmitotic neurons (PMID: 28793256), adult cortical excitatory neurons (PMID: 26087164), and embryonic cortical 7 day ex vivo cultured neurons (PMID: 20393465). Shaded boxes highlight promoters (pink) and gene bodies (blue). IEGs display a bipartite chromatin signature characterized by promoter-H3K27ac and gene body-H3K27me3 in E14.5 barrelette neurons and E15.5 cortical progenitors and postmitotic neurons. In contrast, H3K27me3 is not present on IEG gene bodies in postnatal cortical excitatory neurons, similar to postnatal barrelette neurons (see Fig. 3e and Extended Data Fig. 6b). Also, culturing embryonic cortical neurons for one week results in depletion of the H3K27me3 mark, similar to embryonic hindbrain neuron culture (see Extended Data Fig. 9a). In long genes (for example Nr4a3), H3K27me3 deposition on gene body does not stretch throughout the gene body, but is restricted only to the proximal region downstream of the promoter (also see Extended Data Fig. 3c). b, IEG (Fos, Egr1 and Fosb) genome browser views at E14.5, pre-fixed prior to the dissociation procedure. Shaded boxes highlight promoters (pink) and gene bodies (blue). The presence of the bipartite pattern (H3K27ac + promoter/ H3K27me3+ gene body) in the pre-fixed tissue indicates that it is not induced by the dissociation procedure.
Extended Data Fig. 3 | Genome-wide characterization of the bipartite signature. a and b, Bipartite (a) and bivalent (b) gene rank (x axis) and corresponding rate of correct bipartite and bivalent classification obtained through genome browser visual inspection of individual loci (y axis, fraction of bipartite true-positive, Methods) in E10.5 K20\textsuperscript{tdTomato\++} progenitors and E14.5, E18.5 and P4 Drg11vPrV-ZsGreen/\textsuperscript{+} barrelette neurons. c, Aggregate plot showing profile of H3K27me3 around the transcription start site (TSS) in E14.5 Bip genes (top 100 genes ranked by bipartiteness scores in E14.5 Drg11vPrV-ZsGreen/\textsuperscript{+} barrelette neurons) with long gene length (>10 kb). Promoters (defined as 1 kb upstream to 500 bp downstream of TSS) and gene bodies (from 1 kb to 3 kb downstream of TSS) are highlighted. Note that H3K27me3 on gene bodies does not stretch further than 2-3 kb downstream of the TSS, even when genes are long. d, Genome browser profiles of representative bipartite IEGs (Fos, Egr1) in mouse ESCs. Note that the H3K27me3 mark is deposited not only on downstream (gene body) but also upstream regions of these H3K27ac promoters. e, Scatterplots showing promoter H3K27ac (x axis) and gene body H3K27me3 (y axis) signals in E14.5 rhombomere 3 (r3)-derived K20\textsuperscript{tdTomato\++} hindbrain cells. E14.5 Bip genes identified in Drg11vPrV-ZsGreen/\textsuperscript{+} barrelette neurons are mapped (red dots). Dashed lines indicate thresholds corresponding to a 5% false discovery rate (FDR) based on a gaussian mixture model with two components (for foreground and background, see Methods). Barrelette neuron E14.5 Bip genes show high levels of promoter H3K27ac and gene body H3K27me3 indicating that they are bipartite also in r3-derived K20\textsuperscript{tdTomato\++} hindbrain cells. f, CpG average observed/expected (o/e) ratios in a 100 bp window in E14.5 Bip and E14.5 Biv gene loci. Bins overlapping with the promoter, TSS, and gene body are indicated. g, Transcription factor binding motifs specifically enriched in E14.5 Bip as compared to E14.5 Biv promoters (Methods).
Extended Data Fig. 4 | Coexistence of H3K27ac and H3K27me3 at promoter and gene body of bipartite genes. a, Scheme of sequential ChIP-seq protocol (Supplementary Methods). b and c, Scatterplots comparing H3K27ac (x axis) and H3K27me3 (y axis) signals detected in regions from −1 kb to +3 kb around the transcription start site (TSS) by single ChIP-seq performed with large (2-3 kb) chromatin fragments in E14.5 hindbrain tissue. The colors indicate the corresponding H3K27me3/H3K27ac sequential-ChIP-seq signals, either for all autosomal genes (b), or only for E14.5Bip genes (c). Dashed lines indicate thresholds corresponding to a 5% FDR based on a gaussian mixture model with two components (for foreground and background, see Methods). Stratified by one single ChIP signal (for example H3K27ac), the sequential ChIP signal still correlates with the other (for example H3K27me3), which indicates that single chromatin fragments have been double-marked and thus have been enriched at both steps of the sequential ChIP experiments. d, Genome browser view of bipartite Egr1 gene locus displaying chromatin accessibility (ATAC-seq), 2-3kb-fragment H3K27ac, H3K27me3 and H3K27me3/H3K27ac sequential-ChIP-seq in E14.5 hindbrain tissue. H3K27ac and H3K27me3 coexist on gene body and promoter, respectively. e, Violin plots displaying promoter H3K27ac (left), bulk mRNA-seq (middle, Smart-seq2), and single cell fraction with detected mRNA-seq (right, 10X genomics) of E10.5 K20tdTomato+/− progenitors. E10.5 bipartite (E10.5Bip) genes (green, n = 99 genes, see Methods) and E10.5 non-bipartite genes with Bip-matching promoter H3K27ac levels (blue, n = 99 genes) are compared. E10.5Bip gene transcripts are only detected in as little as 6% of single cells on average. Plots extend from the data minima to the maxima with the white dot indicating median, the box showing the interquartile range and whiskers extending to the most extreme data point within 1.5X the interquartile range. P values are from two-sided Wilcoxon’s tests. NS: not significant (P > 0.05).
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | t-SNE visualization of mouse genes according to chromatin pattern. 

a, Aggregate plot of chromatin profiles (ATAC-seq, ChIP-seq) around transcript start sites (TSSs) of all autosomal genes in E14.5 Drg11vPrV-ZsGreen/+ barrelette neurons. Promoter and gene body regions are highlighted (Methods).

b–d, Two-dimensional (2D) projection on a E14.5, E18.5 and P4-combined t-SNE map of autosomal genes according to chromatin accessibility, H3K27me3, H3K4me2, and H3K27ac levels at promoters and gene bodies in Drg11vPrV-ZsGreen/+ barrelette neurons (Methods).

b, E14.5 (red), E18.5 (blue) and P4 (green) genes are indicated (Supplementary Note).

c, Color-coded t-SNE gene maps according to promoter (top row) and gene body (bottom row) chromatin profiles (columns).

d, Color-coded t-SNE gene maps indicating mRNA levels. Numbers 1–3: example genes in E14.5 barrelette neurons.

e, Chromatin profiles of the example genes in d, namely H3K4me2+/H3K27ac-/H3K27me3-/ATAC- (active, Actb), Polycomb-dependent H3K4me2+/H3K27ac-/H3K27me3+/ATAC+ (permissive, Dlx5) and H3K4me2-/H3K27ac-/H3K27me3-/ATAC- (repressed, Olfr67).

f, color-coded E14.5, E18.5 and P4-combined t-SNE gene maps displaying bipartiteness (left) or bivalency (right) scores in E14.5, E18.5 and P4 barrelette neurons (Methods).

g, E14.5, E18.5 and P4-combined t-SNE map with contour lines indicating regions enriched with bipartite (green) and bivalent (red) genes (Methods).

h, E14.5, E18.5 and P4-combined t-SNE map in which all E14.5 genes are colored according to their developmental change of chromatin state from E14.5 to P4 (Supplementary Note and Methods). i, E14.5Bip genes are subdivided into two subgroups according to their localization on the t-SNE plot (n = 57, E14.5Bip-a genes, orange dots; n = 43, E14.5Bip-b genes, blue dots).

j, Violin plots showing promoter H3K27ac (left), gene body H3K27me3 (middle) and mRNA (right) levels in E14.5Bip-a (orange) and E14.5Bip-b genes (blue) in E14.5 barrelette neurons (Supplementary Note and Methods). Plots extend from the data minima to the maxima with the white dot indicating median, the box showing the interquartile range and whiskers extending to the most extreme data point within 1.5X the interquartile range. P values are from two-sided Wilcoxon's tests.

k and l, E10.5 K20frDsRed/+ progenitor t-SNE maps with bipartiteness or bivalency scores (k) and contour lines (l) at E10.5 (Methods).
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Developmental dynamics of the bipartite chromatin signature. **a**, Developmental dynamics of chromatin profiles (ATAC-seq and ChIP-seq signals in promoter and gene body regions) and mRNA levels of E14.5Bip genes in E14.5, E18.5 and P4 Drg11vPrV-ZsGreen/+ barrelette neurons. Log2 fold changes are calculated with reference to E14.5. At P4, 20% of E14.5Bip genes become expressed (Exp, that is RPKM ≥ 3 at P4), 25% become bivalent (Biv, red dots in Fig. 3d), and the rest (55%) remain bipartite (Bip), (blue, red and green lines, respectively). Bottom: summary diagram. **b**, Genome browser view of the Egr1 locus at the E10.5, E14.5, E18.5 and P4 stages. **c**, 3D interaction map (4C-seq) using the Fos promoter (top left), enhancer 2 (e2, bottom left) and enhancer 5 (e5, bottom right) as viewpoints in E10.5, E14.5, E18.5 and P4 hindbrain tissue. Normalized read per 4C fragment is visualized. **d**, Genome browser views of Fos and Egr1 at the E14.5, E18.5 and P4 stages. **e**, Violin plots showing transcription end site (TES, Methods) RNAPII-S7P (left), RNAPII-S2P (middle), and H3K36me3 (right) levels of E14.5Bip genes at E14.5 and P4 (see **a**). E14.5Bip genes that become expressed (Exp, blue, n = 25) at P4 displayed significantly higher levels of RNAPII-S7P, -S2P and H3K36me3 marks as compared to E14.5Bip genes that become bivalent (Biv, red, n = 20) or remain bipartite (Bip, green, n = 55) at P4: violin deviations between groups are to be compared within the same time point since the time points reflect different batches. Plots extend from the data minima to the maxima with the white dot indicating median, the box showing the interquartile range and whiskers extending to the most extreme data point within 1.5X the interquartile range. P values are from two-sided Wilcoxon’s tests. NS: not significant (P > 0.05).
Extended Data Fig. 7 | See next page for caption.
Extended Data Fig. 7 | Polycomb marking of gene bodies inhibits productive mRNA elongation. a, Spliced transcript expression (transcripts per million, TPM, Methods) of E14.5Bip genes (red) and control genes (blue) with matching distributions of spliced transcripts (Methods) (n = 3 biologically independent replicates). b, Unspliced over total transcript fractions for each gene set in a (Methods). Note the larger fraction of unspliced transcripts of E14.5Bip genes compared to control genes. c, Fractions of spliced over total transcripts of E14.5Bip genes (red) in E14.5 control K20<sup>KO</sup>/+ (WT) and Ezh2<sup>cKO</sup>/+ hindbrain cells (n = 3, biologically independent littermates). d, Violin plots comparing the TSS/whole gene ratios of total RNA-seq reads between Ezh2<sup>cKO</sup>/+ and WT hindbrain cells (see Fig. 5c). e, Scatter plot (left) comparing H3K36me3 levels on the coding region between E14.5 wild-type Hoxa2<sup>WT</sup> (WT) and Ezh2<sup>cKO</sup>/+ hindbrain cells, indicating H3K36me3-positive and negative E14.5Bip genes in red and gray, respectively (Methods). Increased H3K36me3 levels of H3K36me3-positive E14.5Bip genes in Ezh2-depleted cells are further illustrated (right panel). Bars indicate the median. f, (left, middle) Genes with non-zero expression in EedKO and wild-type (WT), and carrying H3K27me3 on gene bodies in mouse ESCs (n = 3457) were subdivided into genes that show up-regulated (red, n = 1067), down-regulated (blue, n = 93) and unchanged (gray, n = 2297) levels of mRNA in EedKO compared with WT ESCs (Supplementary Note, Methods). (right) Violin plots showing log2 fold changes of transcription end site (TES) RNAPII-S2P levels in EedKO mutant compared with WT ESCs. g, Chromatin profiles of Fos in WT and EedKO ESCs. While stalled RNAPII-S5P showed decrease in the promoter (pink highlight), elongating RNAPII-S2P was increased (green highlight) in EedKO ESCs. h, MA-plots comparing mRNA levels of bipartite (Bip) genes (n = 100, red) between full Ezh1KO;Ezh2KO (left) or Ezh2 catalytically inactive Ezh1KO;Ezh2<sup>Y726D</sup> (right) with WT ESCs. d and f, Plots extend from the data minima to the maxima with the white dot or middle bar indicating median, the box showing the interquartile range and whiskers extending to the most extreme data point within 1.5X the interquartile range. P value is from a two-sided Wilcoxon’s test between the two groups.
Extended Data Fig. 8 | IEG bipartite chromatin is necessary to prevent precocious activity-dependent neuronal maturation. 

**a**, Short time (two days) E12.5 ex vivo-cultured Drg11tdTomato/+ hindbrains. After over-night (o/n) treatment with a cocktail of neuronal activity blockers (TDN cocktail = TTX + D-AP5 + NBQX, inhibitors of sodium channel, NMDAR and AMPAR), cultured neurons were treated by 55 mM KCl for 1 hour. Drg11-positive immature trigeminal neurons were FACS-isolated for ATAC-seq analysis. Violin plots visualize log2 fold changes of enhancer chromatin accessibilities in 1 hour KCl-treated neurons as compared to non-treated control neurons. Increased accessibility is selectively detected in KCl-treated neurons at activity-dependent Fos-binding enhancers that normally become open only at P4 (green, n = 85 enhancers) (purple, all non-Fos-binding enhancers that gain accessibilities only at P4, n = 3882 enhancers). Plots extend from the data minima to the maxima with the white dot indicating median, the box showing the interquartile range and whiskers extending to the most extreme data point within 1.5X the interquartile range from the box. P value is from a two-sided Wilcoxon’s test.

**b**, Scatterplots comparing enhancer accessibilities (ATAC) in E14.5 Ezh2 heterozygous control (ctrl) and homozygous mutant (Ezh2cKOHB-RFP) hindbrain cells. All the barrelette enhancers in Drg11vPrV-ZsGreen/+ barrelette neurons (left), non-Fos-binding enhancers that gain accessibilities at P4 as compared with E14.5 Drg11vPrV-ZsGreen/+ barrelette neurons (n = 3882 enhancers, middle, purple), neuronal activity-dependent Fos-binding enhancers that gain accessibilities at P4 as compared with E14.5 Drg11vPrV-ZsGreen/+ barrelette neurons (n = 85 enhancers, left, green) are shown (Methods). 85 activity-dependent Fos-binding enhancers show precocious opening upon H3K27me3 removal at E14.5. Also see Fig. 5d.
Extended Data Fig. 9 | dCas9-UTX overexpression in E12.5 short-term ex vivo cultured neurons. a, H3K27me3 profiles of the Fos locus in E14.5 Dgo1<sup>Pna</sup>-<sup>2Gos</sup> barrelette neurons and E12.5 cultured hindbrain neurons at day 1 and day 7 of culture. One week hindbrain neuron culture results in the loss of the H3K27me3 mark from the Fos gene body, similarly to one-week embryonic cortical neuron culture (Extended Data Fig. 2a); in contrast, short-term (day 1) cultured hindbrain neurons still maintain H3K27me3 levels comparable to E14.5 barrelette neurons. b, H3K27me3 levels at the Fos locus in short-term cultured E12.5 hindbrain neurons overexpressing control dCas9 (green) or dCas9-UTX (red) targeted to Fos gene body: three biological replicates overlaid. Genomic regions targeted by guide-RNAs (gRNAs) are indicated. c, Averaged H3K27me3 profiles of Fos (left) and the rest of the E14.5Bip genes (right). Overexpression of control dCas9 (green) or dCas9-UTX (red) are compared. d, (left) MA-plot comparing H3K27me3 levels of dCas9-UTX against dCas9 targeting to Fos locus. Fos (red dot) shows a loss of gene body H3K27me3 compared to control genes carrying similar levels of H3K27me3 (green dots). (right) Density plot (green line) shows the distribution of the logFC values of the selected control genes (green dots), highlighting the Fos logFC (red line) in the 1.48 % lower tail of the density plot and logFC of E14.5Bip genes (yellow line), indicating slight but significant decrease in Fos H3K27me3, n = 3 biologically independent neuron cultures. e, mRNA levels of Actb, Gapdh and Fos determined by RT-qPCR in short-term cultured E12.5 hindbrain neurons overexpressing control dCas9 (green) or dCas9-UTX (purple) targeted to Actb (left) or Gapdh (right) loci (n = 6 biologically independent neuron cultures). The median expression is indicated by bars. P values are from Welch’s two-sample two-sided t-tests. NS: not significant (P > 0.05).
Extended Data Fig. 10 | Polycomb marking of bipartite gene bodies regulates the rapidity and amplitude of transcriptional response to relevant stimuli.

**a.** Genome browser view of bivalent *Junb* in E14.5 Drg11vPrV-ZsGreen/+ barrelette neurons. Chromatin accessibility (ATAC), H3K4me2, H3K27ac and H3K27me3 are shown.

**b.** mRNA levels of bipartite (*Fos, Egr1*) and bivalent (*Junb*) ARGs, determined by RT-qPCR in E14.5 hindbrain cells treated by KCl for 8 (left) or 30 min (right) (*n* = 4, biologically independent embryos). The median expression is indicated by bars. *P* values are from Welch’s two-sample two-sided *t*-tests. NS: not significant (*p* > 0.05).

**c.** mRNA levels of *Fos* (left) and *Egr1* (right), determined by RT-qPCR in serum-starved WT (green) and EedKO (purple) mouse ESCs (*n* = 4, biologically independent cultured cells) treated with a low (1%) or high (10%) concentration of Fetal Calf Serum (FCS) for 16 minutes. WT ESCs *Fos* and *Egr1* could be induced only after prolonged exposure (that is 16 minutes) to 1% FCS; also see the effects of shorter (that is 8 minutes) time exposure to 1% FCS in Fig. 5f and g. The median expression is indicated by bars. *P* values are from Welch’s two-sample two-sided *t*-tests.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| Item | Confirmed |
|------|-----------|
| The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement | X |
| A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly | X |
| The statistical test(s) used AND whether they are one- or two-sided | X |
| A description of all covariates tested | X |
| A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons | X |
| A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) | X |
| For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted | X |
| Give P values as exact values whenever suitable. | X |
| For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings | X |
| For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes | X |
| Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated | X |

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection | Illumina NextSeq 500, Illumina HiSeq 2500, Zeiss LSM700, StepOnePlus Real-Time PCR System |
|----------------|--------------------------------------------------------------------------------------------|
| Data analysis  | cutadapt (version 1.12, 1.15 and 1.18), STAR (version 2.5.2b), bowtie2 (version 2.3.0, 2.3.3.1 and 2.3.4.2), bowtie (version 1.0.0), Samtools (version 1.2, 1.6 and 1.9), MACS2 (version 2.1.1.20160309 and 2.1.2), salmon (version 0.12.0), R (version 3.5.1 and 3.5.3, version 3.6.1 for single cell RNA-seq, and version 3.6.3 for motif enrichment analysis), TxDb.MmMMusculus.UCSC.mm10.knownGene package (version 3.4.0), QuasR package (version 1.22.1), tximport package (version 1.10.1), rtracklayer package (version 1.42.2), limma package (version 3.38.3), Rtsne package (version 0.13), swissknife package (version 0.10), biomaRt package (version 2.38.0), CellRanger (version 3.0.2), DropletUtils package (version 1.4.3), scater package (version 1.12.2), scran package (version 1.12.1), MASS package (version 7.3-51.1), Reference genome GRCm38/mm10, ENCODE (vM19 transcriptome, vM20 for the single cell RNA-seq), wiggleplotr package (version 1.34.0), UpsetR package (version 1.3.3), Biostats package (version 2.50.2), monaLisa package (version 0.1.28), JASPAR2018, Homer (version 4.10.4), vioplot package (version 0.3.0), Imaris9.2.1, Matlab2019, Fiji/ImageJ2.1.0, GenomicRanges (version 1.34.0) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequencing raw data and processed data used for this study are deposited to ArrayExpress and will be released to the public without restrictions.
mRNA-seq (Smart-seq2), E-MTAB-8314
total RNA-seq (Solo RNA-seq), E-MTAB-8311
ChIP-seq (ChIPmentation), E-MTAB-8317
ATAC-seq, E-MTAB-8313
4C-seq, E-MTAB-8295
single cell RNA-seq (10X Genomics), E-MTAB-8312.

In addition, we used following public datasets.
Mouse cortical culture (GSE21161, GSE60192), mouse embryonic forebrain (GSE93011, GSE52386), mouse adult cortical excitatory neuron (GSE63137), mouse embryonic stem cells (GSE36114, GSE94250), mouse embryonic stem cells for PRC2-KO experiments in Lavaroni et al., (GSE116603), mouse E14.5 heart tissues (GSE82764, GSE82637, GSE82640, GSE78441, ENCSR068YGC), mouse E14.5 liver tissues (GSE78422, GSE82407, GSE82615, GSE82620, ENCSR032HKE) and E10.5 mouse neural crest cells isolated from the frontl nasal process (FNP) (GSE89437).

Furthermore, we used following public databases.
GENCODE v19, M20 (gene annotation), UCSC (mm10 reference genome assembly, gene annotation), JASPAR 2018 (vertebrate transcription factor motifs)

Field-specific reporting
Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design
All studies must disclose on these points even when the disclosure is negative.

Sample size
We chose samples sizes regarding replicates based on the standards of the field (e.g. the ENCODE Data standards, ENCSR534HMF, ENCSR068YGC). We did not perform statistical tests to predetermine samples sizes.

Data exclusions
No data exclusion.

Replication
As for the qPCR and histological experiments, all the experiments are based on at least three biologically independent replicates (animals/cell cultures). As for sequencing experiments, most of experiments have 2-3 biologically independent replicates, with some exceptions (please see the deposited data). Reproducibility among replicates were very high. Choices of replicates of sequencing data does not affect conclusions of each analysis at all, and normally we used averages of each biological replicates to make figures.

Randomization
As for mouse experiments, each biological replicate were prepared from independent litter mates. Set of mice with the same age and sex were raised in the same conditions and randomly allocated to the different experimental groups. As for cell cultures, treatment conditions were randomly allocated to the different experimental groups.

Blinding
No blinding was performed. This does not affect our findings because our experiments using cells and animals were performed in the identical condition and our comparisons were not based on subjective measurements.

Reporting for specific materials, systems and methods
We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants
- Clinical data
- n/a

Methods
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging
- n/a

Antibodies
- Rabbit polyclonal anti-H3K4me2 antibody, Millipore 07-030
- Rabbit polyclonal anti-H3K27me3 antibody, Millipore 07-449
- Rabbit polyclonal anti-H3K27ac antibody, Abcam ab4729
- n/a
Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) ESCs are derived from the lab of Dr. A Wutz. ESCs are originally described in Schoeftner et al., EMBO 2006 (PMID: 16763550)

Authentication Genotypes were confirmed by PCR, western blotting, mRNA-seq and ChIP-seq.

Mycoplasma contamination Cell lines tested negative for Mycoplasma.

Commonly misidentified lines (See ICLAC register) No commonly misidentified lines are used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals Mice were maintained on a mixed C57BL/6J CD1 background and housed in an environmentally controlled room, under a 12-h light:12-h dark cycle. Ambient temperature is 22 +/- 2 °C and humidity is maintained 45-65%. 2-6 months old mice were crossed to generate embryos. Mouse embryos were collected and staged according to Theiler, considering the noon of the day on which vaginal plugs were detected after timed mating as E0.5. Embryos with different sexes were pooled to carry out molecular analysis including sequencing, but only autosomal genes were considered for the further analysis.

Wild animals No involvement of wild animals.

Field-collected samples No involvement of field-collected samples.

Ethics oversight All animal experimental procedures were performed in accordance with Guide for Care and Use of Laboratory Animals, and were approved by the Veterinary Department of the Canton of Basel-Stadt.

Note that full information on the approval of the study protocol must also be provided in the manuscript.
Data deposition

☑ Confirm that both raw and final processed data have been deposited in a public database such as GEO.
☑ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

mRNA (Smart-seq2): Username: Reviewer_E-MTAB-8314, Password: otqfzj0t
total RNA-seq (Solo RNA-seq): Username: Reviewer_E-MTAB-8311, Password: QQwkjnuk
ChIP-seq (Chipmentation): Username: Reviewer_E-MTAB-8317, Password: uvAu66Mq
ATAC-seq: Username: Reviewer_E-MTAB-8313, Password: 9byihgtm
4C-seq: Username: Reviewer_E-MTAB-8295, Password: phgztidc
single cell RNA-seq (10XGenomics): Username: Reviewer_E-MTAB-8312, Password: sjjdcjvt

Files in database submission

E10_K20TMT_H3K27Ac_1 fastq+bigwig
E10_K20TMT_H3K27Ac_2 fastq+bigwig
E10_K20TMT_H3K27Me3_1 fastq+bigwig
E10_K20TMT_H3K27Me3_2 fastq+bigwig
E10_K20TMT_H3K4Me2_1 fastq+bigwig
E10_K20TMT_H3K4Me2_2 fastq+bigwig
E10_K20TMT_H3K4Me2_3 fastq+bigwig
E14_vPrV_H3K27Ac_1 fastq+bigwig
E14_vPrV_H3K27Ac_2 fastq+bigwig
E14_vPrV_H3K27Me3_1 fastq+bigwig+peaks
E14_vPrV_H3K27Me3_2 fastq+bigwig+peaks
E18_vPrV_H3K27Ac_1 fastq+bigwig
E18_vPrV_H3K27Ac_2 fastq+bigwig
E18_vPrV_H3K27Me3_1 fastq+bigwig
E18_vPrV_H3K27Me3_2 fastq+bigwig
E18_vPrV_H3K4Me2_1 fastq+bigwig
E18_vPrV_H3K4Me2_2 fastq+bigwig
P4_vPrV_H3K27Ac_1 fastq+bigwig
P4_vPrV_H3K27Ac_2 fastq+bigwig
P4_vPrV_H3K27Me3_1 fastq+bigwig
P4_vPrV_H3K27Me3_2 fastq+bigwig
P4_vPrV_H3K4Me2_1 fastq+bigwig
P4_vPrV_H3K4Me2_2 fastq+bigwig
E12_1day_H3K27me3 fastq+bigwig
E12_1day_H3K27me3 fastq+bigwig
E12_7day_H3K27me3 fastq+bigwig
E12_ctl_H3K27me3_1 fastq+bigwig
E12_ctl_H3K27me3_2 fastq+bigwig
E12_ctl_H3K27ac fastq+bigwig
E12_TSA_H3K27me3_1 fastq+bigwig
E12_TSA_H3K27me3_2 fastq+bigwig
E12_TSA_H3K27ac fastq+bigwig
E14_H3K27Ac-2K fastq+bigwig
E14_H3K27Me3-2K fastq+bigwig
E14_H3K27Me3AcDouble_1 fastq+bigwig
E14_H3K27Me3AcDouble_2 fastq+bigwig
E14_pCreb fastq+bigwig
P4_pCreb fastq+bigwig
E14_prvPreFix_H3K27Ac fastq+bigwig
E14_prvPreFix_H3K27Me3 fastq+bigwig
E14_H3K36Me3 fastq+bigwig
P4_H3K36Me3_1 fastq+bigwig
P4_H3K36Me3_2 fastq+bigwig
P4_H3K36Me3_3 fastq+bigwig
ESC_wt_H3K4me2 fastq+bigwig
ESC_wt_H3K27ac_1 fastq+bigwig
ESC_wt_H3K27ac_2 fastq+bigwig
ESC_wt_H3K27me3 fastq+bigwig
ESC_wt_Pol2_S2P_1 fastq+bigwig
ESC_wt_Pol2_S2P_2 fastq+bigwig
ESC_wt_Pol2_S5P fastq+bigwig
ESC_EedKO_Pol2_S2P_1 fastq+bigwig
ESC_EedKO_Pol2_S2P_2 fastq+bigwig
ESC_EedKO_Pol2_S5P fastq+bigwig
E12_dCas9_Fos_H3K27Me3_1 fastq+bigwig
E12_dCas9_Fos_H3K27Me3_2 fastq+bigwig
E12_dCas9_Fos_H3K27Me3_3 fastq+bigwig
E12_dCas9Utx_Fos_H3K27Me3_1 fastq+bigwig
E12_dCas9Uttx-Fos_H3K27Me3_3 fastq+bigwig
E12_dCas9Uttx-Fos_H3K27Me3_3 fastq+bigwig
E14_a2CreEzh2homo_H3K36me3 fastq
E14_a2CreEzh2wt_H3K36me3 fastq
E14_K20TMT_H3.3_1 fastq
E14_K20TMT_H3.3_2 fastq
E14_K20TMT_H3K27Ac_1 fastq
E14_K20TMT_H3K27Ac_2 fastq
E14_K20TMT_H3K27Me3_1 fastq
E14_K20TMT_H3K27Me3_2 fastq
E14_K20Ezh2KO_H3K27Ac_1 fastq
E14_K20Ezh2KO_H3K27Ac_2 fastq
E14_K20Ezh2KO_H3K27Me3_1 fastq
E14_K20Ezh2KO_H3K27Me3_2 fastq
E14_Pol2_8WG16 fastq
E14_Pol2_S2P fastq+bigwig
E14_Pol2_S5P fastq+bigwig
E14_Pol2_S7P fastq+bigwig
E18_Pol2_S2P fastq+bigwig
E18_Pol2_S7P fastq+bigwig
P4_Pol2_S2P fastq+bigwig
P4_Pol2_S7P fastq+bigwig
E14_Ring1b fastq
E14_Cdk9 fastq
ESC_WT_NELFB_1 fastq
ESC_WT_NELFB_2 fastq
ESC_EedKO_NELFB_1 fastq
ESC_EedKO_NELFB_2 fastq
E18_pCreb fastq+bigwig
E18_Jmd3wt_K27Me3_1 fastq
E18_Jmd3wt_K27Me3_2 fastq
E18_Jmd3KO_K27Me3_1 fastq
E18_Jmd3KO_K27Me3_2 fastq

Genome browser session (e.g. UCSC)
bigwig files used for figures are deposited in ArrayExpress.

Methodology

Replicates
We chose samples sizes based on previous experiments and literature surveys. Most of samples have 2-3 replicates, with some exceptions.

Sequencing depth
Libraries were sequenced by Illumina HiSeq 2500 machine (50bp read length, single end). Typically total sequencing depths were 30-60M reads, and more than 90% reads with high quality mapping to mm10.

Antibodies
- Rabbit polyclonal anti-H3K4me2 antibody, Millipore 07-030
- Rabbit polyclonal anti-H3K27me3 antibody, Millipore 07-449
- Rabbit polyclonal anti-H3K27ac antibody, Abcam ab4729
- Rabbit polyclonal anti-H3.3 antibody, Millipore 09-838
- Mouse monoclonal anti-RNAPII 8WG16 antibody, Covance MMS-126R
- Mouse monoclonal anti-Ser5P RNAPII 4H8 antibody, Covance MMS-128P
- Rat monoclonal anti-Ser7P RNAPII 4E12 antibody, Millipore 04-1570
- Rat monoclonal anti-Ser2P RNAPII 3E10 antibody, Active Motif 61083
- Rabbit polyclonal anti-H3K36me3 antibody, Abcam ab9050
- Rabbit polyclonal anti-Ser133P CREB antibody, Millipore 17-10131
- Rabbit monoclonal anti-Ring1b antibody, Cell Signaling 5694
- Rabbit monoclonal anti-Cdk9 antibody, Abcam ab239364
- Rabbit polyclonal anti-NELF-B antibody, Abcam ab237027
- Rabbit monoclonal anti-H3K27me3 antibody, Cell Signaling 9733

Peak calling parameters
A hidden semi-markov model (HSMM) was used to identify Polycomb regions (chromatin that is positive for the H3K27me3 histone modification) genome-wide. This approach was used instead of typical peak callers since the size of Polycomb regions varies a lot across the genome, from short peaks up to large domains. A hidden markov model (HMM) or peak callers like MACS2 will be unable to find optimal parameters that model both short and long regions and typically do not account for this variation. We therefore tiled the genome into 500 bp regions and the H3K27me3 ChIP-seq read count per tile was log2-transformed with a pseudo-count of 1. In an HMM setting these are the observations, and the hidden states that are to be inferred are “Polycomb” and “non-Polycomb”. The HSMM in addition models the duration of a state and thus was better suited to call Polycomb regions.

The mhsmm (version 0.4.16) R package was used. The hsmmspec function was used to estimate the model parameters assuming a Gaussian distribution on the emissions (the log2-count data of the two states) and a gamma distribution on the sojourn (duration of a state). The hsmmfit function then used the model’s estimated parameters with mstep=mstep.norm and maxit=100, to call Polycomb and non-Polycomb regions.

Data quality
All samples were quality controlled by qQCReport function of QuasR package version 1.22.1. GC bias was also assessed and normalized according to necessity.
Software

cutadapt (version 1.15 and 1.18), bowtie2 (version 2.3.3.1 and 2.3.4.2), Samtools (version 1.6 and 1.9), bowtie (version 1.0.0), R (version 3.5.1 and 3.5.3), QuasR package (version 1.22.1), TxDb.MmuscUs.UCSC.mm10.knownGene package (version 3.4.0), limma package (version 3.38.3), Rtsne package (version 0.13), swissknife package (version 0.10), mclust package (version 5.4.1 and 5.4.3), mhsmm package (version 0.4.16), MASS package (version 7.3-51.1), vioplot package (version 0.3.0), MACS2 (version 2.1.2) (to define background peaks used in the normalization process), wiggleplotr package (version 1.6.1)

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Dissected tissues were kept in PBS 1X on ice, then treated with papain digestion mix (papain 10mg/ml/ cysteine 2.5mM/ HEPES pH7.4 10mM/ EDTA 0.5mM/ DMEM 0.9X) for 3-5 minutes at 37°C and immediately put on ice. Tissue was rinsed by ice-cold DMEM 1X, and dissociated by pipetting and filtered. Cells were collected by FACS. Processing of these cells was adapted for further analyses (e.g. RNA-seq, ATAC-seq and ChIP-seq). During optimization of the dissociation protocol, we confirmed that our short-term enzymatic treatment does not cause the artificial expression of immediate early genes, by comparing Actinomycin D treated and un-treated conditions.

Instrument

BD FACS Aria III

Software

FlowJo LLC

Cell population abundance

Signal positive fractions were always determined by comparing against equivalent tissues from negative control embryos that do not express fluorescent markers. Negative controls were prepared from same litter mates. Resorting ensured >97 % purity.

Gating strategy

Cells were preliminary selected by scatter areas (SSC-A vs SSC-A). Doublet was discriminated by firstly assessing forward scatter (FSC-W vs FSC-H) and next by assessing side scatter (SSC-W vs SSC-H). Subsequently cells were sorted based on fluorescent markers of interest.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.