A gene regulatory network for neural induction

Katherine E. Trevers,[1,7] Hui-Chun Lu,[1,7] Youwen Yang,[1,3,8] Alexandre Thiery,[2,8] Anna C. Strobl,[1] Božena Pálinkášová,[1] Nidia M. M. de Oliveira,[1] Irene M. de Almeida,[1] Mohsin A. F. Khan,[1] Natalia Moncaut,[1,4] Nicholas M. Luscombe,[5,6] Leslie Dale,[1] Andrea Streit,[2] Claudio D. Stern[1,9,*]

[1]Department of Cell and Developmental Biology, University College London, London WC1E 6BT, United Kingdom.
[2]Centre for Craniofacial and Regenerative Biology, King's College London, London SE1 9RT, United Kingdom.
[3]Present address: School of Cardiovascular Medicine & Sciences, King's College London, London, SE5 9NU, United Kingdom.
[4]Present address: Cancer Research UK Manchester Institute, The University of Manchester, Alderley Park, SK10 4TG, United Kingdom.
[5] The Francis Crick Institute, London, UK.
[6] UCL Genetics Institute, Department of Genetics, Environment and Evolution, University College London, London, UK
[7]These authors contributed equally.
[8]These authors contributed equally.
[9]Lead contact
*Correspondence: c.stern@ucl.ac.uk

Abstract

During early vertebrate development, signals from a special region of the embryo, the organizer, can re-direct the fate of non-neural ectoderm cells to form a complete, patterned nervous system. This is called neural induction and has generally been imagined as a single signaling event, causing a switch of fate. Here we undertake a comprehensive analysis, in very fine time-course, of the events following exposure of ectoderm to the organizer. Using transcriptomics and epigenomics we generate a Gene Regulatory Network comprising 175 transcriptional regulators and 5,614 predicted interactions between them, with fine temporal dynamics from initial exposure to the signals to expression of mature neural plate markers. Using in situ hybridization, single-cell RNA-sequencing and reporter assays we show that neural induction by a grafted organizer mimics normal neural plate development. The study is accompanied by a comprehensive resource including information about conservation of the predicted enhancers in different vertebrate systems.

Introduction

One of the most influential studies in developmental biology was the discovery, 100 years ago, that a small region of the vertebrate embryo, named the "organizer", can induce ectodermal cells that do not normally contribute to the neural plate to form a complete, patterned nervous system (Spemann, 1921; Spemann and Mangold, 1924). In amphibians, where these experiments were initially conducted, the "organizer" resides in the dorsal lip of the blastopore. A few years later, Waddington demonstrated that an equivalent region exists in birds (duck and chick) and mammals (rabbit) (Waddington, 1933, 1934,

1936, 1937; Waddington and Schmidt, 1933). In these groups, an equivalent organizer is located at the tip of the primitive streak, a structure known as Hensen's node (Hensen, 1876). This interaction between the organizer and the responding ectoderm, which causes the latter to acquire neural plate identity, has been termed "neural induction"(Gallera, 1971b; Gurdon, 1987; Nieuwkoop et al., 1952; Saxen, 1980; Spemann and Mangold, 1924; Stern, 2005; Storey et al., 1992).

Neural induction has often been imagined as a single event, "switching" the fate of the responding tissue from non-neural to neural. But it is clear from timed transplantation experiments that the responding ectoderm requires several hours' (about 12) exposure to the organizer (Gallera, 1971b; Gallera and Ivanov, 1964; Streit et al., 1998) to acquire its new identity in a stable way ("commitment"). Recent work has also revealed that the expression of many genes changes over time after grafting an organizer, suggesting that the process has considerable complexity (Albazerchi and Stern, 2007; Gibson et al., 2011; Papanayotou et al., 2013; Papanayotou et al., 2008; Pinho et al., 2011; Sheng et al., 2003; Stern, 2005; Streit et al., 2000; Streit et al., 1998; Streit et al., 1997; Streit and Stern, 1999; Trevers et al., 2018). What happens during this period? Is it possible to define distinct steps, and perhaps the molecular events that represent "induction" and "commitment"? Surprisingly, given the long time since the discovery of neural induction, these questions have hardly been addressed. To begin to answer them requires a comprehensive approach to identify and model the key interactions between genes and transcriptional regulators inside the responding cells that accompany the responses to signals from the organizer over time.

Grafting an organizer to a site that does not normally contribute to neural tissue but is competent to do so provides the opportunity to study the progression of neural induction relative to "time-0", the moment when the organizer is first presented to the tissue. This allows for the separation of neural inductive events from other processes that occur during normal neural plate development (for example mesendoderm formation and patterning), because sites that are remote enough from the normal neural plate and are competent to respond to an organizer graft can generate a complete neural tube without induction of mesoderm (Dias and Schoenwolf, 1990; Hornbruch et al., 1979; Storey et al., 1992).

Here we have taken advantage of these properties, together with recent major technological advances in transcriptomics and epigenetic analysis, to undertake a comprehensive dissection of the molecular events that accompany neural induction in fine time-course and to generate the first detailed Gene Regulatory Network (GRN) for this process. The GRN comprises 175 transcriptional regulators and 5,614 predicted interactions between them over the course of neural induction. We then compare the spatial and temporal properties of these changes with development of the normal neural plate of the embryo using *in situ* hybridization, single cell RNA-sequencing and reporter assays to test the activity of some key gene regulatory elements. We present a comprehensive resource allowing visualization and querying of all of these interactions and regulatory elements on a genome-wide level. Together, our study offers a comprehensive view of the genetic hierarchy of neural induction and normal neural plate development over time.

Results

Transcriptional profiling identifies responses to neural induction in time course

A graft of Hensen's node to an extraembryonic region of competent non-neural ectoderm, the inner area opaca (Dias and Schoenwolf, 1990; Gallera and Ivanov, 1964; Storey et al., 1992; Streit et al., 1998), induces a mature, patterned neural tube after 15h of culture (Figure 1A). To characterize the events over this period, the induction of several neural markers was assessed. *OTX2* is first expressed in the pre-primitive-streak stage epiblast at EGKXII-XIII (Albazerchi and Stern, 2007; Pinho et al., 2011) and is induced by 7/7 grafts after 3h (Figures 1B-F). *SOX2* is first expressed at HH4+/5 (Rex et al., 1997; Sheng et al., 2003; Streit et al., 1998; Streit et al., 1997) and requires 9h for induction (8/8) (Figures 1G-K). *SOX1* is expressed weakly in the neural tube as it forms at around HH7-8 (Stavridis et al., 2010; Uchikawa et al., 2011) and is induced by 50% of grafts after 12h (Figures 1L-P). These time courses confirm that a sequence of events occurs in response to grafted nodes over 0-12h (Figure S1A), culminating in the acquisition of neural plate/tube identity (Albazerchi and Stern, 2007; Gibson et al., 2011; Papanayotou et al., 2013; Papanayotou et al., 2008; Pinho et al., 2011; Sheng et al., 2003; Stern, 2005; Streit et al., 2000; Streit et al., 1998; Streit et al., 1997; Streit and Stern, 1999; Trevers et al., 2018).

To uncover the full transcriptional responses to neural induction over this period, we performed RNAseq at three time points following Hensen's node grafts: 5h (to identify early "pre-neural" responses), 9h (when *SOX2* expression defines neural specification) and 12h, as cells start to express *SOX1*. Hensen's node was grafted from HH4- chick donors to a region of competent epiblast (inner area opaca) of chick hosts at the same stage. Embryos were cultured for the desired period of time, after which the graft was removed and the underlying "Induced" tissue was collected, as well as control ("Uninduced") ectoderm from the corresponding region on the contralateral side of the same embryos (Figure 1Q). Competent area opaca at HH4-, representing a "0h" starting point for the time course, was also sampled. Transcripts were sequenced and mapped to the chicken genome. When comparing "Induced" and "Uninduced" reads at each time point, DESeq analyses identified 8673 differentially expressed transcripts based on a ±1.2 $\log_2$ fold change (Figure 1R-T). Of these, 4130 were upregulated (enriched in "induced" tissue) and 4543 were downregulated (depleted in "induced" tissue) relative to the "uninduced" counterpart.

To construct a GRN we focused on 482 transcription factors and chromatin modifiers that are differentially expressed within these data. Of these "transcriptional regulators", 255 are upregulated, 203 are downregulated and 24 have more complex expression dynamics (Figures 1U and S1B). Grouping transcripts based on the timing of their response (Figure 1V-W) reveals that 120 are differentially expressed throughout (53 upregulated and 67 downregulated), while others are unique to particular time points. Therefore, neural induction involves changes in the expression of many genes over a relatively short period, and is therefore complex and dynamic.

Due to this unexpected complexity, we chose to increase the time resolution of this analysis. NanoString nCounter was used to quantify the gene expression changes of transcriptional regulators at 6 time points (1, 3, 5, 7, 9 and 12h after the node graft) in "induced" tissue compared to "uninduced" ectoderm. By consolidating the data from RNAseq and NanoString, a set of refined expression profiles was established for 156 transcriptional regulators that are enriched and 57 that are depleted in induced tissues (Table S1-4). These represent the core components of our GRN.

Epigenetic changes identify chromatin elements associated with neural induction

We next sought to identify the regions of non-coding chromatin that govern responses to neural induction. Histone modifications regulate gene expression by altering chromatin structure. For example, H3K27ac is associated with actively transcribed genes and their enhancers, whereas transcriptionally inactive regions are often marked by H3K27me3 (Bonn et al., 2012; Creyghton et al., 2010; Heintzman et al., 2009; Kharchenko et al., 2011; Rada-Iglesias et al., 2011; Tiwari et al., 2008; Tolhuis et al., 2011; Zentner et al., 2011). To detect H3K27 chromatin marks that change during neural induction, ChIPseq was performed on induced and uninduced ectoderm following 5, 9 or 12h of a node graft (Figure 2A).

H3K27ac and H3K27me3 enriched regions were identified genome-wide by comparison to a genomic input sample. Chromatin sites were categorized according to their histone signatures across induced and corresponding uninduced tissues at each time point (Figure S2A). Sites that become acetylated and/or demethylated in induced tissues compared to uninduced were considered to undergo "activation" (Indices 1-3); those that become deacetylated and/or methylated undergo "repression" (Indices 4-6). Chromatin marked by both H3K27ac (activation) and H3K27me3 (repression) marks in either tissue were described as "Poised" (Indices 7-12). The remaining indices (13-16) define sites that do not change marks between the two conditions.

Chromatin sites undergoing activation are enriched for H3K27ac marks in induced tissues at each time point (Figure 2B-D). Sites of repression are more varied: those belonging to this category at 5h lose acetylation, whereas a number of sites gain methylation in induced tissues at 9h and 12h. Very few sites are "poised" at any time. Overall, as neural induction progresses, the number of sites undergoing activation increases, along with a reduction in those that do not change state.

Next, we focused on changes associated with the 213 genes encoding transcriptional regulators for which refined expression profiles were established by NanoString analysis (see above). "Constitutive" CTCF-bound sites (putative insulators) flanking these genes were obtained from chicken CTCF-ChIPseq data (Kadota et al., 2017; Khan et al., 2013). They were used to predict the boundaries within which H3K27 marks were identified by ChIPseq; these regions ("loci") are up to 500kb in length and may contain several genes (Figure S2B). Within each locus, the H3K27ac or H3K27me3 enriched peaks were categorized as before (Figure S2A). The 213 selected loci contain a total of 6,971 sites that change in response to a node graft (Indices 1-10; Figure S2A). Our analysis identifies these as putative regulatory elements that are controlled epigenetically during the process.

A Gene Regulatory Network for neural induction.

Having identified transcriptional responses to neural induction and their accompanying chromatin changes, we combined them to construct a GRN to describe the time course of these events and to illustrate predicted interactions between transcriptional regulators. The open source platform BioTapestry (Longabaugh et al., 2005, 2009; Paquette et al., 2016) is now used extensively to provide a visually intuitive representation of developmental networks (Arnone and Davidson, 1997; Betancur et al., 2010; Davidson, 1990; Davidson et al., 1998; Simoes-Costa and Bronner, 2015; Thiery et al., 2020). Importantly, this software allows visualization of dynamic changes in the interactions between regulators and their target genes.

A BioTapestry model for regulatory gene interactions

A custom computational pipeline was developed to integrate these complex time course data (Figure 3A). The 213 transcriptional regulators were selected as candidate components of the network based on their responses to a node graft in RNAseq and NanoString, and associated chromatin changes according to ChIPseq. We aimed to generate a GRN for interactions occurring during neural induction. The interactions were predicted by screening of transcription factor binding sites within the putative regulatory sites with indices 1-3 and 7 (Figure S2C-E) associated with a member of the GRN.

Positive or negative regulations were then modeled between regulators and their putative targets by overlaying the predicted transcription factor binding profiles with consolidated expression profiles from RNAseq and NanoString using the regulatory rules shown in Figure S3A-B. At each time point, changes in expression of a putative regulator that could lead to the same change in a candidate target are modeled as positive interactions. Negative, or inhibitory, interactions are predicted when changes in expression of the regulator could cause the opposite change in a downstream target. Genes lacking both regulatory inputs and outputs with other genes in the network were excluded.

The resulting GRN for neural induction, depicting interactions between 175 genes, was represented using BioTapestry (Figure 3B and Suppl File S1). Most targets are predicted to receive multiple positive and/or negative inputs, which are initiated at, and can act across, multiple time points. For example, the non-neural gene *GATA2* is predicted to be repressed by *TFAP2C* after just 1h of a node graft, followed by increasing repression via *HIF1A*, *MEIS2* and numerous other factors over the remaining time course. On the other hand, *BLIMP1* is predicted to be induced by *TFAP2C* after 1h, alongside *SNAI1*, *ETV1*, *ETV4*, *MGA* and *SOX4* (Suppl File S2). In total, 5,614 interactions are predicted to occur between these components, highlighting the intricate and highly dynamic sequence of regulatory events accompanying neural induction.

Incorporating individual regulatory sites and their dynamics into the network

BioTapestry networks usually represent each gene once, with multiple regulatory inputs, and each regulator as a single input into the target gene. However, it is now known that gene expression is controlled by multiple regulatory elements, each with characteristic spatial and temporal activity. In the past, such elements were identified by constructing many reporter fragments by nested deletions, whose activity was then tested *in vivo*; for example, the pioneering work of H. Kondoh's lab in identifying many enhancers controlling *SOX2* expression (Okamoto et al., 2015; Uchikawa et al., 2003). Our epigenetic analysis reveals that in almost all cases there are indeed multiple putative regulatory sites associated with genes that change expression, which can be in different chromatin states (Figure 4A). By predicting the changes that occur as sites undergo epigenetic "activation" or "repression" (using Indices 1-7; Figure S3A), we can model the different contributions of these chromatin elements. Since BioTapestry lacks a notation for distinct regulatory elements of individual genes, as an example, we generated a sub-network showing 6 candidate *cis*-regulatory regions of *BRD8* at 5, 9 and 12h as separate targets, where each target is represented with the BioTapestry notation used for genes (Figure 4B-D). This reveals which specific elements contain binding sites for each regulating transcription factor. One element (site1 in Figure 4B) is initially active, however it does not receive inputs from other GRN

components. Site1 is then methylated at 9h and 12h as an increasing number of other elements become active, presumably to stabilize and maintain *BRD8* expression. Although it is not practical to generate a full BioTapestry representation of the entire network including all the elements in their different states in this way, all identified elements and their predicted activity are presented in the associated UCSC Browser tracks (https://genome.ucsc.edu/s/stern_lab/Neural_Induction_2021) together with the predicted binding sites relevant to the network contained within each site; a full list is also given in Table S5-7.

Neural induction by a graft of Hensen's node recapitulates normal neural plate development

To what extent is the induction of an ectopic neural plate by a graft comparable to the events of normal neural plate development? To address this, we used three complementary approaches. First, we explored the spatial and temporal patterns of expression of many components of the network during normal embryo development. Then we conducted single cell RNAseq analysis of stages of neural plate development concentrating on these network components and their temporal hierarchical organization. Finally, we tested the activity of several of the identified putative regulatory elements in normal embryos to assess whether their activity in the neural plate *in vivo* resembles the hierarchy revealed by node grafts.

Spatiotemporal expression of GRN components during normal neural plate development

A previous study identified genes whose expression differs between EGKXII-XIII (pre-primitive-streak) epiblast, neural plate at HH6-7 and non-neural ectoderm (Trevers et al., 2018). When these embryonic tissues (Figure 5A) are compared to neural induction, we find that the pre-streak epiblast is most similar to induced ectoderm after 5h of a node graft, while 9h and 12h time points are more closely related to mature neural plate (Figure 5B). This encouraged us to generate a detailed spatiotemporal map of transcripts during embryonic neural plate development. Genes that are differentially expressed in induced tissues (using a ±1.2 $\log_2$ fold change compared to uninduced in RNAseq, p-value < 0.05) were selected. *In situ* hybridization was performed for genes encoding 174 transcriptional regulators (including 123 that are represented in the GRN) at 4 stages: pre-streak (EGKXII-XIII), primitive streak (HH3-4), head process/early neural plate (HH5-6) and neural fold/tube (HH7-9).

There is a striking correspondence between neural induction by a node graft and neural development in terms of the spatiotemporal patterns observed. In addition to confirming the expression of known neural markers (*ERNI*, *SOX3*, *SOX2*, and *SOX1*; Figure 5C-F), the predictions from node grafts uncover many novel responses such as *CITED4* (*ERNI*-like), *MAFA* (*SOX3*-like), *GLI2* and *ZNF423* (*SOX2*-like) (Figure 5G-J). In total, 84/89 transcriptional regulators that are induced by a node graft were detected in neural tissues at some stage, 79 of which had not previously been associated with neural induction (Figure 5 and Suppl File S3). Only 5 genes (*HEY1*, *RFX3*, *STOX1*, *TAF1A* and *ZIC2*) could not be detected by *in situ* hybridization - these are induced at very low FPKM levels according to RNAseq. Likewise, a high proportion (65/85; 76%) of transcripts that are downregulated by a node graft are depleted in neural tissues relative to non-neural territories. This includes known non-neural markers such as *GATA2* and *MSX2* and the extraembryonic marker *HIF2A* (Figure 5K, L, O). Neural/non-neural border markers (such as *DLX5*, *MSX1* and *TFAP2C*; Figures 5M and Suppl File S3) were also downregulated, supporting

suggestions that neural induction involves the repression of an early border identity (Trevers et al., 2018). We also uncover several other genes (such as *ISX*, *GRHL3* and *VGLL1*) that are excluded from neural tissues (Figure 5P-R).

Importantly, this approach reveals quite subtle temporal changes in expression. For example, *CITED4* is sharply upregulated after 5h of a node graft but is later downregulated (Figure 5G). In the embryo, it is expressed in pre-streak epiblast and prospective neural plate at HH3-4, but these transcripts are then excluded from the neural plate at HH5-6. *CDX4* and *HOXA2* are initially repressed by a node before their expression levels increase after 12h – just as they become enriched in the posterior neural tube from HH7-9 (Figure 5N and Suppl File S3). These observations confirm that transcriptional responses to a node graft closely follow the events that occur during development of the embryonic neural plate.

The GRN describes a temporal hierarchy occurring in single cells during neural development

To quantify the expression of multiple genes during normal neural plate development and to compare this to the temporal hierarchy of our GRN, we assessed the expression of network components in individual cells at different stages of normal neural development. Single cell RNAseq (scRNAseq) was performed on ectoderm dissected from chick embryos at HH4, HH6, HH8 and HH9+. These broad regions included prospective or mature neural plate and neural tube as well as non-neural ectoderm and neural plate border (Figure 6A).

The non-linear dimensional reduction method UMAP was used to visualize unbiased cell clustering in low dimensionality space (Figure 6B). Cells collected from the two earliest developmental stages (HH4 and HH6) each form a distinct group, whereas cells from later stages (HH8 and HH9+) are clustered primarily according to cell type (Figure 6C). Cell identities were assigned to the 15 clusters using well-established markers (Figure S4A). This revealed clusters corresponding to placodal (cluster 8), neural crest (clusters 10, 11), and neural tube identities (clusters 4, 6, 7, 9, 13, 14) within the HH8 and HH9+ cell populations. At HH6, neural plate cells expressing *LIN28B*, *SETD2* and *SOX2* (cluster 0) are distinct from non-neural plate progenitors (cluster 3), which express *GATA2*, *DLX5* and *TFAP2A*. At HH4, prospective neural cells express *MAFA*, *ING5* and *LIN28B* (clusters 1, 2 and 5) whereas cluster 12 expresses the node marker ADMP.

To identify markers that are co-expressed in a cell-type specific manner, gene module analysis was performed using Antler (Delile et al., 2019). For a module to be retained for further consideration, we applied the criterion that at least 50% of its genes must be differentially expressed (logFC > 0.25, FDR < 0.001) in at least one cell cluster relative to the rest of the dataset. This condition defines seventeen distinct modules (Figure S4B), which support our cell type classifications. Module-3 shows co-expression of *PAX2*, *WNT4* and *NKX6-2* and is upregulated in prospective midbrain clusters 9 (at HH8) and 6 (at HH9+). Module-10 is defined by co-expression of *PAX7*, *SOX8*, *SOX10* and *TFAP2B* and is associated with neural crest (clusters 10, 11). Module-25 has low levels of expression of *GATA2* and *ID3* (clusters 1, 2, 5, 12) at HH4, which are then further downregulated at later stages, except in non-neural cells at HH6 (cluster 3) and placodal cells at HH8 (cluster 8).

To validate the temporal hierarchy of GRN components, their expression was assessed in prospective neural plate and neural tube cell populations (clusters 0, 1, 2, 4, 5, 6, 7, 9, 13, 14). Within this “neural” population, principal component (PC) analysis revealed that PC1 inversely correlates with developmental stage (Figure S4C), therefore cells were subsequently ranked according to their position

along PC1 to explore expression dynamics. To visualize the expression of GRN components, genes were grouped according to the time point at which they are first expressed after a node graft; their average (normalized and scaled) expression was then plotted (Figure 6D). This reveals that components upregulated within 1, 3 or 5h of a node graft are initially highly expressed in prospective neural cells and that their expression decreases progressively across PC1. GRN components that are induced after 7, 9 or 12h are expressed at higher levels later, corresponding to maturing neural plate and neural tube. Taken together, our single cell RNAseq analysis uncovers the changes in gene expression accompanying development of the normal neural plate. Within this, GRN components defined from node graft experiments are appropriately expressed, suggesting that the temporal hierarchy of the neural induction assay closely replicates normal neural plate development.

Responses to neural induction reveal enhancers that drive neural gene expression during normal development

Do the regulatory elements predicted by the GRN generated from Hensen's node grafts correspond to those that regulate normal development of the neural plate in the embryo? To help define the putative enhancers more precisely, we combined the histone profiling described in the previous sections with ATACseq (Buenrostro et al., 2013; Buenrostro et al., 2015; Corces et al., 2017) to identify regions of open/active chromatin from induced and uninduced tissues after 5, 9 and 12h after a node graft. These tissues were then compared to ChIPseq (for H3K27ac/me3) and ATACseq performed on pre-primitive-streak epiblast and mature neural plate at HH6-7.

We chose 6 putative enhancers of GRN components for more detailed exploration. These were identified within the previously described CTCF loci using the overlapping combination of H3K27ac marks and ATACseq reads to predict enhancers that are likely to be active (Figures S5-7). Each region was cloned into the pTK-EGFP vector (Uchikawa et al., 2003) to drive GFP expression via a minimal promoter; the reporter was then electroporated into the pre-primitive-streak stage (EGKX-XI) epiblast or the prospective neural plate at HH3/3+ (Figure 7A) and embryos cultured for 5h or to the desired stages.

All 6 enhancers drive neural-specific GFP expression in a spatiotemporal pattern strikingly similar to the genes that lie in their proximity (Figure 7B-O and S5-7). For example, SOX11_enh1 and SETD2_enh1 are enriched for H3K27ac in induced tissues after 5h (Figure S5A-B). Both are active in the epiblast at EGKXII-XIII and prospective neural plate at HH3-4, when *SOX11* and *SETD2* are expressed (Figure S5C-AY). SOX2_enh3 is "poised" after 5h of neural induction, but acquires an active signature at 9h when it gains H3K27ac and ATAC reads (Figure S6A). As expected, this enhancer is not active in the prospective neural plate at HH4 (Figure S6C-H); GFP positive cells are detected later when *SOX2* is expressed in the neural plate proper at HH5-6 (Figure S6I-N). Likewise, GLI2_enh1 and SIX3_enh2 also become active after 9h; GLI2_enh1 is detected throughout the neural plate, whereas SIX3_enh2 is restricted to the forming forebrain at HH6 (Figure S7A-Z). These five enhancers remain active in the neural tube in correspondence with H3K27ac marks and ATAC reads that persist after 12h of a node graft.

CDX1_enh1 is also active in the caudal neural tube at HH7-9, when it also weakly labels some presomitic mesoderm cells. Its activity is detected at HH5-6, but only in a few cells of the caudal lateral epiblast adjacent to Hensen's node and not in the neural plate (Figure S6U-AL). This sparse activity overlaps closely with the discrete region from which neuro-mesodermal precursors, which are known to express *CDX1* (Gouti et al., 2017), are thought to contribute to both neural and somite tissues (Brown and

Storey, 2000; Tzouanacou et al., 2009). Since the *CDX1* gene is expressed earlier (HH3-4), but only in mesoderm and primitive streak, the activity of this enhancer is likely to be associated with NMPs.

Taken together, these comparisons between node grafts and normal embryos strongly suggest that histone profiles that change in response to a node graft (upon which our network is based) can predict regulatory elements relevant to neural development. This establishes our GRN as a robust tool to describe the genetic hierarchy associated with the acquisition of neural identity.

Browser resources

In addition to the BioTapestry GRN and other associated data described above, the results of this project, including data from node grafts in time course and from normal neural plate development, are made available on the UCSC Browser at https://genome.ucsc.edu/s/stern_lab/Neural_Induction_2021. This incorporates RNAseq, ChIPseq and ATACseq for all expressed and repressed genes (not limited to GRN components) and the annotated putative regulatory regions with predicted binding sites for GRN components. Finally, to help to predict potential conservation of the regulatory sites across species, we performed DREiVe analysis (Khan et al., 2013; Sosinsky et al., 2007; Streit et al., 2013) for domains within a 500kb window around each of the genes represented in the GRN, using human (hg38), mouse (mm10), rat (rn6), golden eagle (aquChr2), zebrafish (danRer10) and chicken (galGal5). Blocks of sequence with conserved motifs are included as a track in the browser.

Discussion

This study unfolds the full complexity of the responses to signals from the "organizer" in very fine time-course up to the time of appearance of mature neural plate markers, such as *SOX1*. We show that the hierarchy of responses to a graft of Hensen's node into a remote part of the embryo closely recapitulates the events of normal neural plate development, including appropriate activity of enhancers identified based on the graft assay. Our GRN, comprising 175 transcriptional regulators and 5,614 putative regulatory interactions between them, allows prediction of the consequences of manipulating the activity of these components. This represents the most comprehensive example to date of a network modelled on both transcriptional and chromatin changes that occur in the embryo. These changes were selected in an unbiased way by comparing induced and uninduced tissues to identify events specific to neural induction. By following this process over very short intervals (1-3h), we are able to predict interactions likely to occur via direct transcription factor binding events. This is accompanied by a rich resource (deposited in the UCSC genome browser) that includes comprehensive genome-wide datasets in time-course for all transcribed loci, epigenetic marks and annotations for prediction of evolutionary conservation of the identified enhancer elements. This is the first comprehensive description of the progression of molecular events of neural induction in any species, and the first GRN for this process.

Is it possible to identify specific changes associated with the key concepts of developmental biology, such as "induction" and "commitment" (Slack, 1991; Stern, 2004) to a neural fate? These concepts have classically been defined based on specific embryological manipulations. Based on timed organizer grafts followed by their removal, a critical time was established at 12-13 hours after a graft (Gallera, 1965,

1971a, b; Gallera and Ivanov, 1964), using similar conditions to those of our node grafts here. This represents the end-point of our GRN which coincides approximately with the onset of expression of the definitive neural tube marker *SOX1* (Pevny et al., 1998).

Previous functional studies have established that competent ectoderm cannot respond to BMP-inhibitors like Chordin unless it has previously been exposed to a graft of the organizer for at least 5 hours (Streit et al., 1998). This led to the identification of a few genes whose expression is associated with this transition (Gibson et al., 2011; Papanayotou et al., 2013; Pinho et al., 2011; Sheng et al., 2003; Streit et al., 2000), but the present experiments greatly expand and enrich this by placing genes in the context of a comprehensive network of gene interactions. Future experiments can take advantage of this information to determine the mechanisms that cause epiblast cells to become sensitive to BMP inhibition at this time point, and the upstream signals responsible. For the time being we feel that functional definitions of the developmental biology concepts will remain key to defining the important molecular changes that underlie these properties of developing tissues, especially given the complexity of the changes in gene expression accompanying neural induction.

Our analysis reveals thousands of sites across the genome whose marks change with the progression of neural induction. Each locus contains genes whose expression changes during this cascade and several putative regulatory elements. As previously described for other systems such as embryonic stem cells in culture, regulatory elements may be decorated simultaneously by marks associated with activation (e.g. H3K27ac) and with repression (e.g. H3K27me3); these have been called “poised” enhancers (Creyghton et al., 2010; Cui et al., 2009; Heintzman et al., 2009; Xu et al., 2009) and often interpreted to mean that the site is in transition from one state to the other, or perhaps able to go either way. Our fine time course certainly reinforces this concept; in many cases we observe that cells pass from one state (e.g. H3K27me3 without acetylation) to the opposite, passing through a “poised” state at the intermediate time point. This could suggest that these changes can occur sequentially in a cell. However, it is also possible that this progression can reflect cells within a tissue that do not undergo the change synchronously, but some cells start earlier than their neighbors, the tissue gradually becoming more homogeneous. This sort of progression is sometimes seen by in situ hybridization or time-lapse movies of reporter activity (for example to visualize the activity of enhancer N2 associated with the SOX2 gene (Papanayotou et al., 2008; Uchikawa et al., 2003). Most likely these “poised” sites reflect a combination of these two reasons.

Among the many regulatory elements predicted by our analysis are some that have been described previously (Iida et al., 2020; Okamoto et al., 2015; Uchikawa et al., 2003; Uchikawa et al., 2011). Their analysis was based on nested deletions followed by testing the activity of the fragments in a vector containing a heterologous minimal promoter. In our study we initially predicted the elements based on a combination of marks and regions marked by ATACseq as being “open”. In some cases, there are differences in the length of the sequence considered to act as an enhancer by both studies. For example, our SOX2_enh3 partially overlaps with, and its behavior upon electroporation mimics the activity of, the previously identified *SOX2* D1 enhancer (Iida et al., 2020; Okamoto et al., 2015). Therefore, our ChIPseq analysis can verify known enhancers as well as identifying many new candidates.

In conclusion, our study provides the first GRN to represent the changes that accompany the process of neural induction after grafting an organizer to an ectopic site, and how it relates to normal neural plate development, in fine time course. We also provide a comprehensive resource including not only the network components (transcriptional regulators with predicted cross-regulatory interactions), but also a

genome-wide view of all transcriptional changes, H3-K27 acetylation/methylation and chromatin conformation (ATACseq) during this process, from the time cells are initially exposed to neural inducing signals to the time at which embryological experiments suggest that the tissue has become "committed" to a neural identity. We envisage that the GRN and the accompanying resource will be invaluable tools to study other important aspects of early neural development, such as the signaling inputs responsible for these changes, the molecular basis of the competence of cells to respond to neural inducing signals at different locations and stages, and aspects of neural patterning leading to regional diversification of the central nervous system. To aid in cross-species comparisons and possible studies on evolutionary changes that have occurred in neural induction, the browser resource also includes the results of computational prediction of conservation of the key GRN loci among several vertebrate species, using DREiVe, a computational tool designed to identify conserved sequence motifs and patterns rather than precise sequence conservation (Khan et al., 2013; Sosinsky et al., 2007; Streit et al., 2013).

Acknowledgements: This study was funded by grants from NIH (R01 MH 60156), MRC (G0400559), Wellcome Trust (063988) and BBSRC (BB/R003432/1 and BB/K007742/1) to CDS and BBSRC (BB/K006207/1) to AS. The work of NML is supported by the Francis Crick Institute which receives its core funding from Cancer Research UK (FC010110), the UK Medical Research Council (FC010110), and the Wellcome Trust (FC010110). The single-cell RNA-seq data analyses were performed using infrastructure from the Crick Scientific Computing science technology platform. NML is a Winton Group Leader in recognition of the Winton Charitable Foundation's support towards the establishment of the Francis Crick Institute.

Main Figure Legends

Figure 1. Transcriptional profiling identifies responses to neural induction in time course
(A) Hensen's node (HN) was grafted from HH3-4 donors to the inner area opaca (yellow) of hosts at the same stage. Ectopic neural tubes expressing *SOX2*, *OTX2* and *HOXB1* are induced after 15h of culture.
(B-F) Expression of neural markers compared to their time course of induction by a node graft: *OTX2*; first expressed in pre-streak epiblast (EGKXII-XIII) and induced by grafts after 3h.
(G-K) *SOX2*; first expressed in the neural plate at HH5-6 and induced after 9h.
(L-P) *SOX1*; first expressed in the forming neural tube at HH7-8 and induced after 12h (3/6 cases).
(Q) Identifying transcriptional responses to a node graft. HN was grafted from HH4- donors to HH4- hosts. The HN graft was removed and underlying "Induced" (IN) and contralateral "Uninduced" (UN) ectoderm isolated after 5, 9 or 12h. Uninduced "0h" ectoderm from HH4- embryos was also dissected. Samples were processed by RNAseq.
(R-T) Induced and corresponding uninduced tissues were compared at each time point to identify differentially expressed transcripts. Volcano plots show upregulated (green) and downregulated (blue) transcripts (> ±1.2 $\log_2$ fold change and <0.05 p-value).
(U-W) Venn diagrams represent 482 genes encoding transcriptional regulators (U) that are upregulated (V) or downregulated (W) at different time points.

Figure 2. Epigenetic changes identify chromatin elements associated with neural induction
(A) Hensen's node (at HH4-) was grafted to hosts of the same stage. Embryos were cultured for 5, 9 or 12h before HN was removed and Induced (IN) and contralateral uninduced (UN) ectoderm collected. ChIPseq was performed for H3K27ac and H3K27me3.

(B-D) Putative regulatory sites were predicted according to the H3K27 profiles of IN and UN tissues at each time point (see Fig S2A). They include sites undergoing "Activation" or "Repression", being "Poised", or showing no change. Heat maps illustrate the enrichment of H3K27ac (blue) and H3K27me3 (purple) in IN and UN tissues within ±2.0kb from the peak center (PC) for each annotated group. Graphs plot the distributions of H3K27ac and H3K27me3 enrichment for each group.

Figure 3. Constructing a BioTapestry model for regulatory gene interactions
(A) Computational pipeline to combine transcriptomic and epigenetic time course data, to build a GRN for neural induction. The output data are available to view in the UCSC browser.
(B) A GRN comprising 5,614 predicted interactions between 175 components is visualized using BioTapestry.

Figure 4. A subnetwork incorporating individual regulatory sites and their dynamics
(A) UCSC browser view of RNAseq and ChIPseq tracks associated with *BRD8*. *BRD8* is upregulated after 9h and 12h of neural induction (green bars). The *BRD8* regulatory locus (grey box; chr13:10028086-10091079) is defined by flanking CTCF-bound sites. Seven putative regulatory sites (S1-7) within this domain were predicted based on the ChIPseq H3K27 profiles. This includes sites that undergo activation (indices 1-3, 7, colored in dark blue), repression (indices 4-6, colored in cyan) or show no change (colored in orange). Transcription factor binding sites by network components are shown; green for components that are expressed at the same time point and blue for those that are not. A BioTapestry subnetwork was generated from these predicted binding sites and expression profiles. Site 3, active at 12h, is not shown in the subnetwork as there is no TF that is expressed at 12h and predicted to bind to it.
(B) *BRD8* regulation during neural induction: *BRD8* is initially not expressed; site 1 is active but is not predicted to be bound by other GRN components.
(C) *BRD8* is upregulated after 9h; sites 2 and 6 undergo activation and could be bound by various transcription factors that are also expressed. Site 1 undergoes repression.
(D) *BRD8* expression is maintained at 12h; regulators potentially bind to sites 2, 4, 5 and 7. Site 6 is no longer predicted to be active.

Figure 5. A spatiotemporal map of the normal embryo for transcripts differentially expressed during neural induction
(A) Central epiblast from pre-primitive-streak stage embryos (PS_cEpi) at EGKXII-XIII, neural plate (NP) at HH6-7 and corresponding non-neural ectoderm were dissected and processed by RNAseq.
(B) Principal component analysis comparing prospective neural and non-neural tissues from the normal embryo to Induced (IN) and Uninduced (UN) ectoderm 5, 9 or 12h after a node graft.
(C-R) *In situ* hybridization of genes encoding transcriptional regulators at 4 stages of embryonic development (EGKXII-XIII, HH3-4, HH5-6 and HH7-9) compared to their RNAseq expression after 5, 9 and 12h of a node graft. Bar charts plotted RNAseq values (FPKM) in Induced and uninduced tissues. Bars are shaded green when genes are upregulated in induced tissue (FPKM >10 in induced tissue and FC > 1.5 compared to uninduced). Bars are shaded blue when genes are downregulated (FPKM >10 in uninduced tissue and FC < 0.5 compared to uninduced). When a gene is neither upregulated nor downregulated, the bars are colored in dark grey in induced and light grey in uninduced. Genes represented within the neural induction GRN are underlined.

Figure 6. The GRN describes a temporal hierarchy occurring in single cells during neural development

(A) Broad regions of prospective neural plate, neural plate border and non-neural ectoderm were dissected at HH4, HH6, HH8 and HH9+ and processed for scRNAseq. *PAX7* expression marks the border between neural and non-neural tissues.
(B) UMAP displaying unbiased clustering of cells into 15 clusters (0-14) representing distinct cell identities.
(C) UMAP displaying the developmental stage from which the cells were collected; cells from HH4 and HH6 form independent clusters, whereas HH8 and HH9+ are more intermixed.
(D) Temporal hierarchy of expression of GRN components in single cells during neural plate development. Prospective neural plate and neural tube cells (clusters 0, 1, 2, 5, 6, 7, 9, 13 and 14) were selected and ranked according to PC1 (see Figure S4C). The average gene expression (mean +/- SE) is plotted for groups of components that are induced at each time point in the GRN, compared to housekeeping genes (HK). A general additive model was fitted to visualize the smoothed expression profile across the ranking of PC1 (using bins of 2.5).

Figure 7. Enhancers identified by the neural induction assay have activity consistent with the expression of the gene during normal neural plate development
(A) Putative enhancers were cloned into the E-pTK-EGFP vector and co-electroporated into a broad region of the pre-streak stage (EGKX-XI) or HH3+ epiblast together with DsRed.
(B-C) Activity of SOX11_enh1 and SETD2_enh1 at EGKXII-XIII;
(D-G) Activity of SOX11_enh1, SETD2_enh1, SOX2_enh3 and CDX1_enh1 at HH4;
(H-K) HH5-6;
(L-O) HH7-9.

References

Albazerchi, A., and Stern, C.D. (2007). A role for the hypoblast (AVE) in the initiation of neural induction, independent of its ability to position the primitive streak. Dev Biol *301*, 489-503.

Arnone, M.I., and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. Development (Cambridge, England) *124*, 1851-1864.

Betancur, P., Bronner-Fraser, M., and Sauka-Spengler, T. (2010). Assembling neural crest regulatory circuits into a gene regulatory network. Annu Rev Cell Dev Biol *26*, 581-603.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczyński, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nat Genet *44*, 148-156.

Brown, J.M., and Storey, K.G. (2000). A region of the vertebrate neural plate in which neighbouring cells can adopt neural or epidermal fates. Curr Biol *10*, 869-872.

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

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol *109*, 21.29.21-21.29.29.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B.*, et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

Cui, K., Zang, C., Roh, T.Y., Schones, D.E., Childs, R.W., Peng, W., and Zhao, K. (2009). Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell *4*, 80-93.

Davidson, E.H. (1990). How embryos work: a comparative view of diverse modes of cell fate specification. Development (Cambridge, England) *108*, 365-389.

Davidson, E.H., Cameron, R.A., and Ransick, A. (1998). Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. Development (Cambridge, England) *125*, 3269-3290.

Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., and Sagner, A. (2019). Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. Development (Cambridge, England) *146*.

Dias, M.S., and Schoenwolf, G.C. (1990). Formation of ectopic neurepithelium in chick blastoderms: age-related capacities for induction and self-differentiation following transplantation of quail Hensen's nodes. Anat Rec *228*, 437-448.

Gallera, J. (1965). Quelle est la durée nécessaire pour déclencher des inductions neurales chez le Poulet? Experientia *21*, 218-219.

Gallera, J. (1971a). Différence de la reactivité à l'inducteur neurogène entre l'ectoblaste de l'aire opaque et celui de l'aire pellucide chez le poulet. Experientia *26*, 1953-1954.

Gallera, J. (1971b). Primary induction in birds. Advances in morphogenesis *9*, 149-180.

Gallera, J., and Ivanov, I. (1964). La compétence neurogéne du feuillet externe du blastoderme de Poulet en fonction du facteur 'temps'. Journal of Embryology and Experimental Morphology *12*, 693-711.

Gibson, A., Robinson, N., Streit, A., Sheng, G., and Stern, C.D. (2011). Regulation of programmed cell death during neural induction in the chick embryo. Submitted.

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2017). A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. Dev Cell *41*, 243-261.e247.

Gurdon, J.B. (1987). Embryonic induction--molecular prospects. Development (Cambridge, England) *99*, 285-306.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W.*, et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Hensen, V. (1876). Beobachtungen über die Befruchtung und Entwicklung des Kaninchens und Meerschweinchens. Z Anat EntwGesch *1*, 353-423.

Hornbruch, A., Summerbell, D., and Wolpert, L. (1979). Somite formation in the early chick embryo following grafts of Hensen's node. J Embryol Exp Morphol *51*, 51-62.

Iida, H., Furukawa, Y., Teramoto, M., Suzuki, H., Takemoto, T., Uchikawa, M., and Kondoh, H. (2020). Sox2 gene regulation via the D1 enhancer in embryonic neural tube and neural crest by the combined action of SOX2 and ZIC2. Genes to Cells *25*, 242-256.

Kadota, M., Hara, Y., Tanaka, K., Takagi, W., Tanegashima, C., Nishimura, O., and Kuraku, S. (2017). CTCF binding landscape in jawless fish with reference to Hox cluster evolution. Sci Rep *7*, 4957-4957.

Khan, M.A., Soto-Jimenez, L.M., Howe, T., Streit, A., Sosinsky, A., and Stern, C.D. (2013). Computational tools and resources for prediction and analysis of gene regulatory regions in the chick genome. Genesis (New York, NY : 2000) *51*, 311-324.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T.*, et al.* (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480-485.

Longabaugh, W.J.R., Davidson, E.H., and Bolouri, H. (2005). Computational representation of developmental genetic regulatory networks. Developmental Biology *283*, 1-16.

Longabaugh, W.J.R., Davidson, E.H., and Bolouri, H. (2009). Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1789*, 363-374.

Nieuwkoop, P.D., Botternenbrood, E.C., Kremer, A., Bloesma, F.F.S.N., Hoessels, E.L.M.J., Meyer, G., and Verheyen, F.J. (1952). Activation and organization of the Central Nervous System in Amphibians. J Exp Zool *120*, 1-108.

Okamoto, R., Uchikawa, M., and Kondoh, H. (2015). Sixteen additional enhancers associated with the chicken Sox2 locus outside the central 50-kb region. Dev Growth Differ *57*, 24-39.

Papanayotou, C., De Almeida, I., Liao, P., Oliveira, N.M.M., Lu, S.-Q., Kougioumtzidou, E., Zhu, L., Shaw, A., Sheng, G., Streit, A.*, et al.* (2013). Calfacilitin is a calcium channel modulator essential for initiation of neural plate development. Nat Commun *4*, 1837.

Papanayotou, C., Mey, A., Birot, A.M., Saka, Y., Boast, S., Smith, J.C., Samarut, J., and Stern, C.D. (2008). A mechanism regulating the onset of Sox2 expression in the embryonic neural plate. PLoS Biol *6*, e2.

Paquette, S., Leinonen, K., and Longabaugh, W. (2016). BioTapestry now provides a web application and improved drawing and layout tools [version 1; peer review: 3 approved]. F1000Research *5*.

Pevny, L.H., Sockanathan, S., Placzek, M., and Lovell-Badge, R. (1998). A role for SOX1 in neural determination. Development (Cambridge, England) *125*, 1967-1978.

Pinho, S., Simonsson, P.R., Trevers, K.E., Stower, M.J., Sherlock, W.T., Khan, M., Streit, A., Sheng, G., and Stern, C.D. (2011). Distinct steps of neural induction revealed by Asterix, Obelix and TrkC, genes induced by different signals from the organizer. PLoS ONE *6*.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

Rex, M., Orme, A., Uwanogho, D., Tointon, K., Wigmore, P.M., Sharpe, P.T., and Scotting, P.J. (1997). Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. Dev Dyn *209*, 323-332.

Saxen, L. (1980). Neural induction: past, present, and future. Curr Top Dev Biol *15 Pt 1*, 409-418.

Sheng, G., dos Reis, M., and Stern, C.D. (2003). Churchill, a zinc finger transcriptional activator, regulates the transition between gastrulation and neurulation. Cell *115*, 603-613.

Simoes-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. Development (Cambridge, England) *142*, 242-257.

Slack, J. (1991). From egg to embryo: Regional specification in early development, second edn (Cambridge: Cambridge University Press).

Sosinsky, A., Honig, B., Mann, R.S., and Califano, A. (2007). Discovering transcriptional regulatory regions in Drosophila by a nonalignment method for phylogenetic footprinting. Proc Natl Acad Sci U S A *104*, 6305-6310.

Spemann, H. (1921). Die Erzeugung thierischer Chimären durch heteroplastische Transplantation zwischen Triton cristatus und taeniatus. Wilh Roux' Arch EntwMech Organ *48*, 533-570.

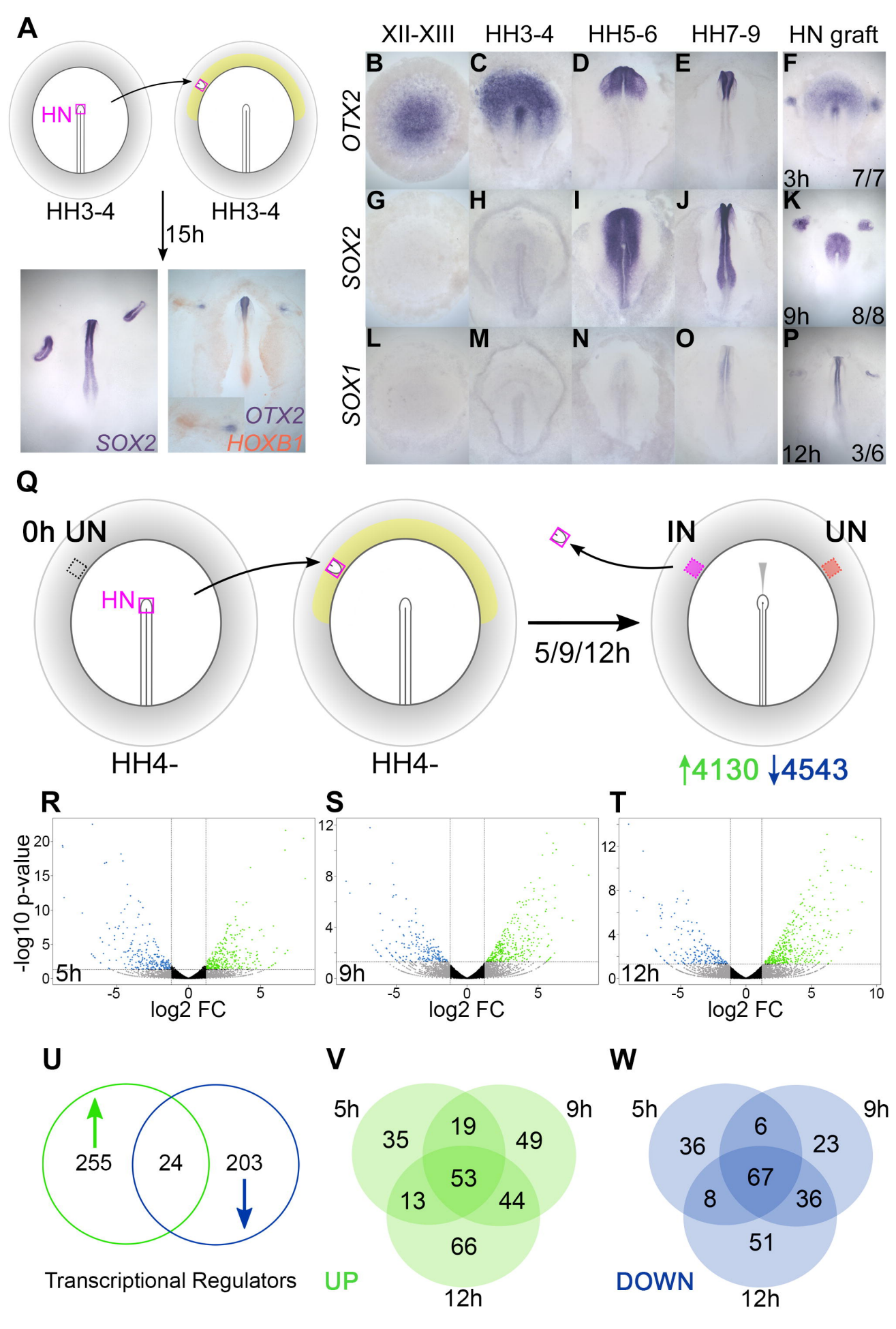
Spemann, H., and Mangold, H. (1924). Über Induktion von Embryonalanlagen durch Implantations artfremder Organisatoren. Roux's Arch EntwMech Org *100*, 599-638.

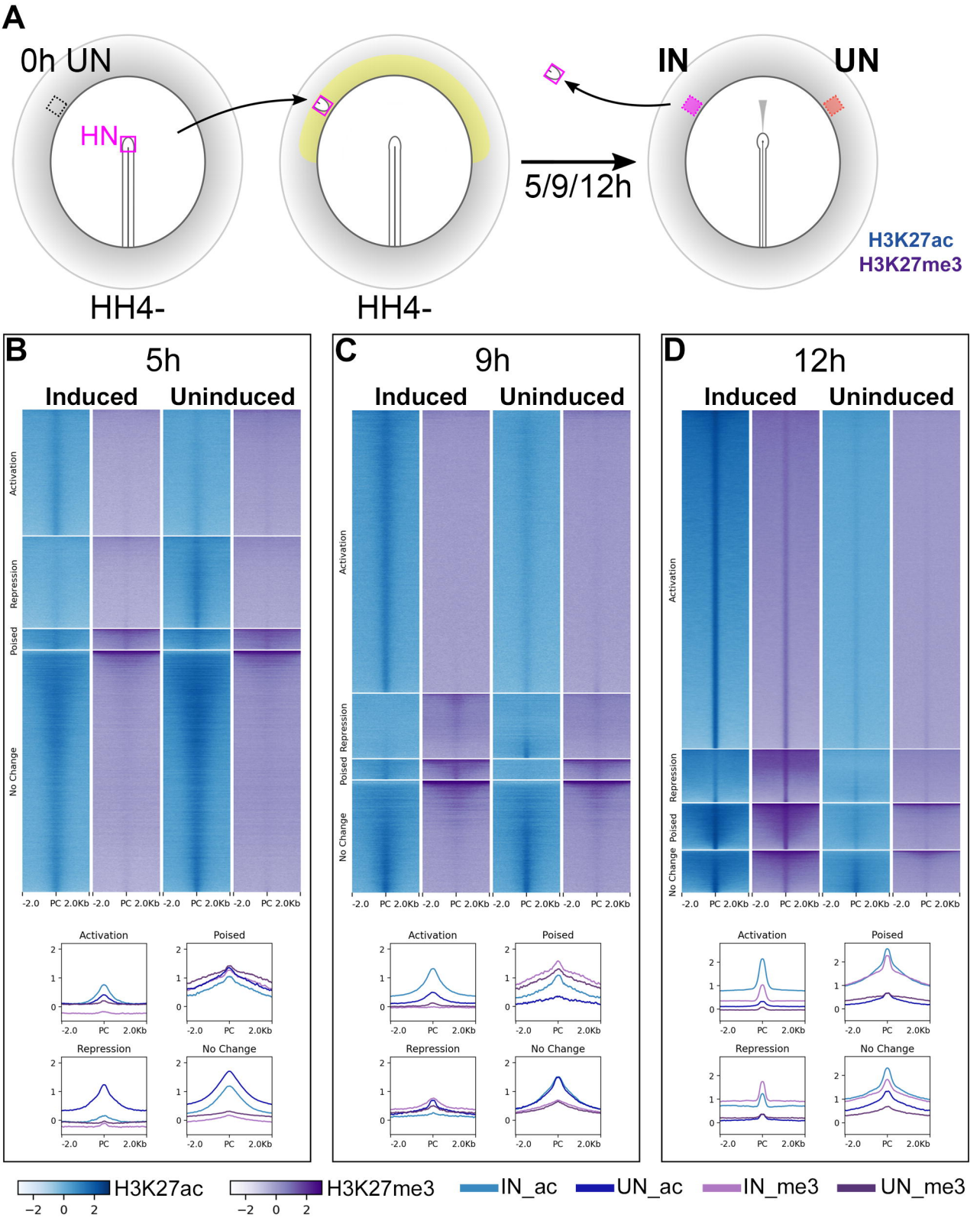
Stavridis, M.P., Collins, B.J., and Storey, K.G. (2010). Retinoic acid orchestrates fibroblast growth factor signalling to drive embryonic stem cell differentiation. Development (Cambridge, England) *137*, 881-890.

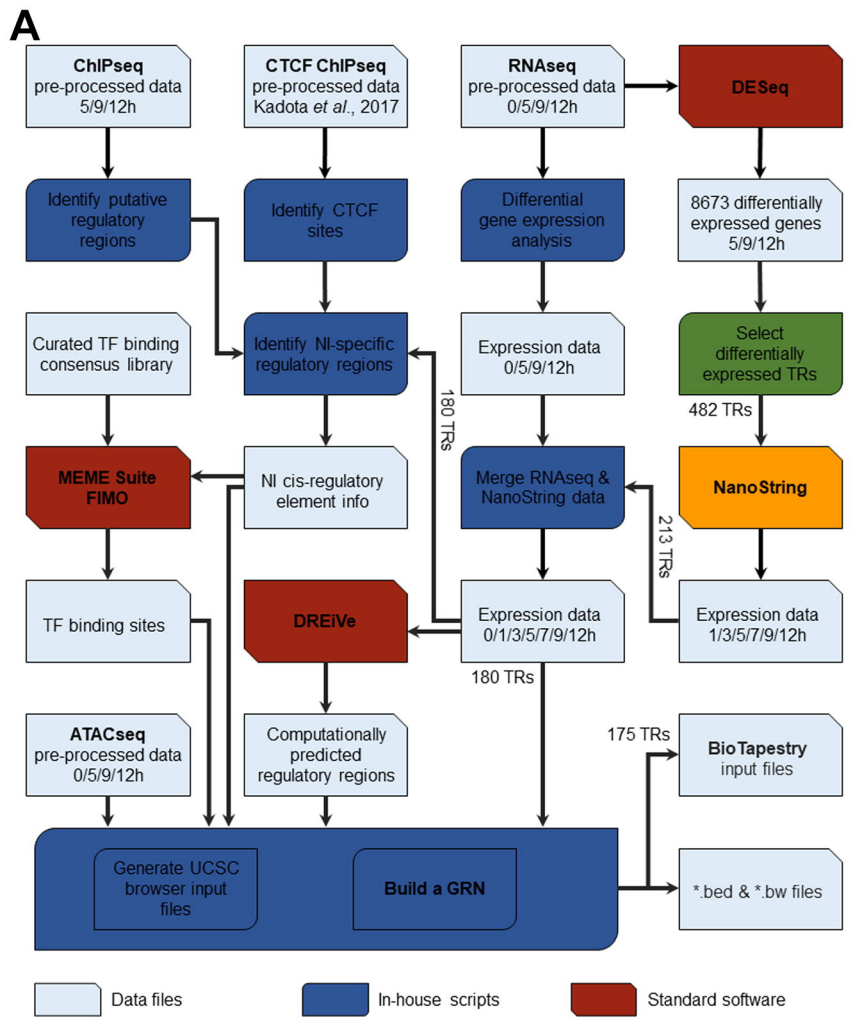
Stern, C.D. (2004). Gastrulation : from cells to embryo (Cold Spring Harbor, N.Y. ; [Great Britain]: Cold Spring Harbor Laboratory Press).
Stern, C.D. (2005). Neural induction: old problem, new findings, yet more questions. Development (Cambridge, England) *132*.
Storey, K.G., Crossley, J.M., De Robertis, E.M., Norris, W.E., and Stern, C.D. (1992). Neural induction and regionalisation in the chick embryo. Development (Cambridge, England) *114*, 729-741.
Streit, A., Berliner, A.J., Papanayotou, C., Sirulnik, A., and Stern, C.D. (2000). Initiation of neural induction by FGF signalling before gastrulation. Nature *406*, 74-78.
Streit, A., Lee, K.J., Woo, I., Roberts, C., Jessell, T.M., and Stern, C.D. (1998). Chordin regulates primitive streak development and the stability of induced neural cells, but is not sufficient for neural induction in the chick embryo. Development (Cambridge, England) *125*, 507-519.
Streit, A., Sockanathan, S., Pérez, L., Rex, M., Scotting, P.J., Sharpe, P.T., Lovell-Badge, R., and Stern, C.D. (1997). Preventing the loss of competence for neural induction: HGF/SF, L5 and Sox-2. Development (Cambridge, England) *124*, 1191-1202.
Streit, A., and Stern, C.D. (1999). Establishment and maintenance of the border of the neural plate in the chick: involvement of FGF and BMP activity. Mechanisms of development *82*, 51-66.
Streit, A., Tambalo, M., Chen, J., Grocott, T., Anwar, M., Sosinsky, A., and Stern, C.D. (2013). Experimental approaches for gene regulatory network construction: the chick as a model system. Genesis (New York, NY : 2000) *51*, 296-310.
Thiery, A., Buzzi, A.L., and Streit, A. (2020). Cell fate decisions during the development of the peripheral nervous system in the vertebrate head. Curr Top Dev Biol *139*, 127-167.
Tiwari, V.K., McGarvey, K.M., Licchesi, J.D., Ohm, J.E., Herman, J.G., Schübeler, D., and Baylin, S.B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. PLoS Biol *6*, 2911-2927.
Tolhuis, B., Blom, M., Kerkhoven, R.M., Pagie, L., Teunissen, H., Nieuwland, M., Simonis, M., de Laat, W., van Lohuizen, M., and van Steensel, B. (2011). Interactions among Polycomb domains are guided by chromosome architecture. PLoS Genet *7*, e1001343.
Trevers, K.E., Prajapati, R.S., Hintze, M., Stower, M.J., Strobl, A.C., Tambalo, M., Ranganathan, R., Moncaut, N., Khan, M.A.F., Stern, C.D.*, et al.* (2018). Neural induction by the node and placode induction by head mesoderm share an initial state resembling neural plate border and ES cells. Proc Natl Acad Sci U S A *115*, 355-360.
Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., and Nicolas, J.F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev Cell *17*, 365-376.
Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y., and Kondoh, H. (2003). Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. Dev Cell *4*, 509-519.
Uchikawa, M., Yoshida, M., Iwafuchi-Doi, M., Matsuda, K., Ishida, Y., Takemoto, T., and Kondoh, H. (2011). B1 and B2 Sox gene expression during neural plate development in chicken and mouse embryos: Universal versus species-dependent features. Development, Growth & Differentiation *53*, 761-771.
Waddington, C.H. (1933). Induction by the primitive streak and its derivatives in the chick. Journal of Experimental Biology *10*, 38-48.
Waddington, C.H. (1934). Experiments on embryonic induction. J exp Biol *11*, 211-227.
Waddington, C.H. (1936). Organizers in Mammalian Development. Nature *138*, 125.
Waddington, C.H. (1937). Experiments on determination in the rabbit embryo. Arch Biol *48*, 273-290.
Waddington, C.H., and Schmidt, G.A. (1933). Induction by heteroplastic grafts of the primitive streak in birds. Zeitschr Wiss Biol Abt D Roux Arch Ent Vncklungsmech Organ *128*, 522-563.

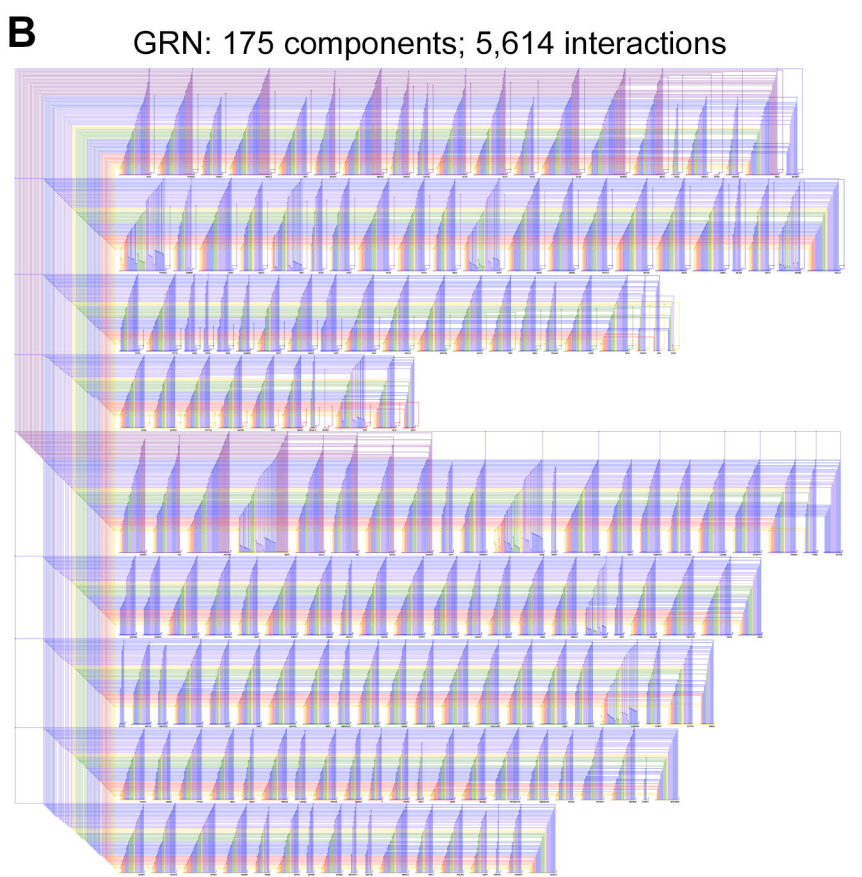
Xu, J., Watts, J.A., Pope, S.D., Gadue, P., Kamps, M., Plath, K., Zaret, K.S., and Smale, S.T. (2009). Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. Genes Dev *23*, 2824-2838.

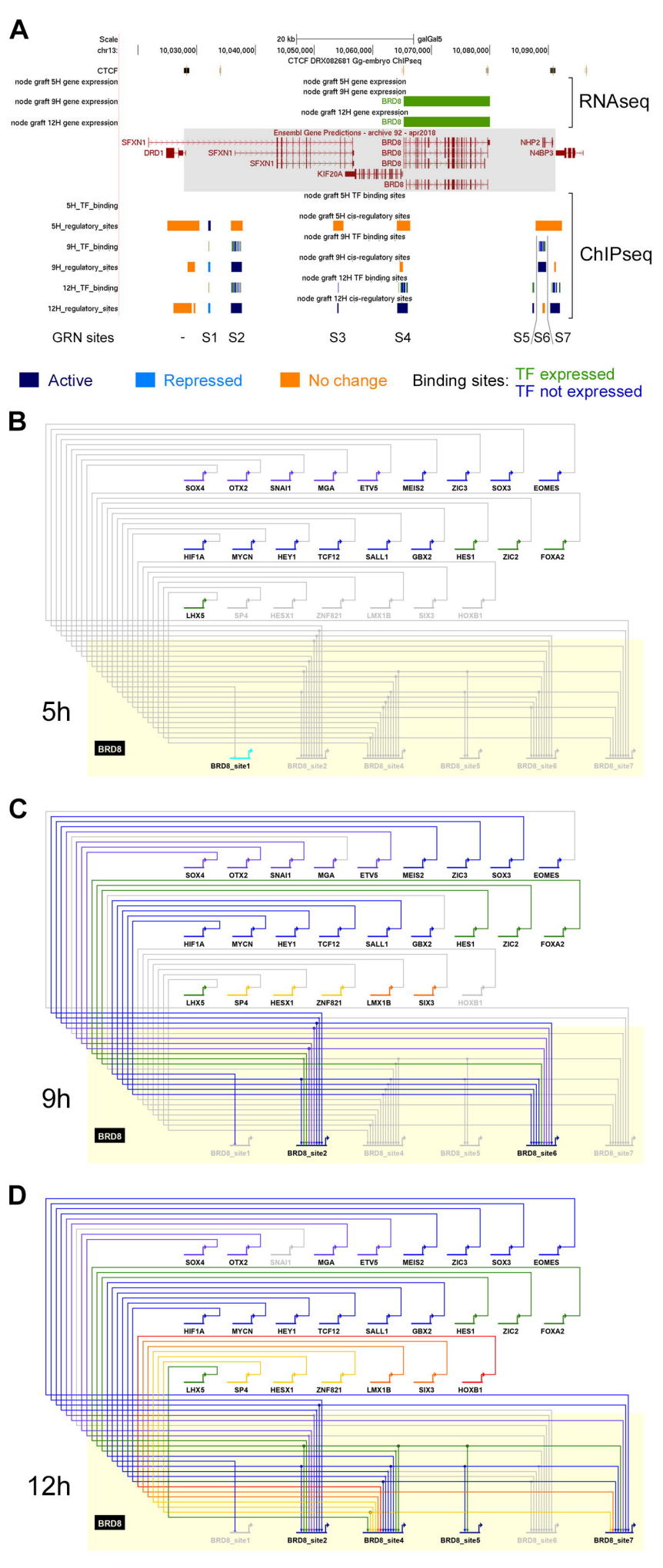
Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res *21*, 1273-1283.

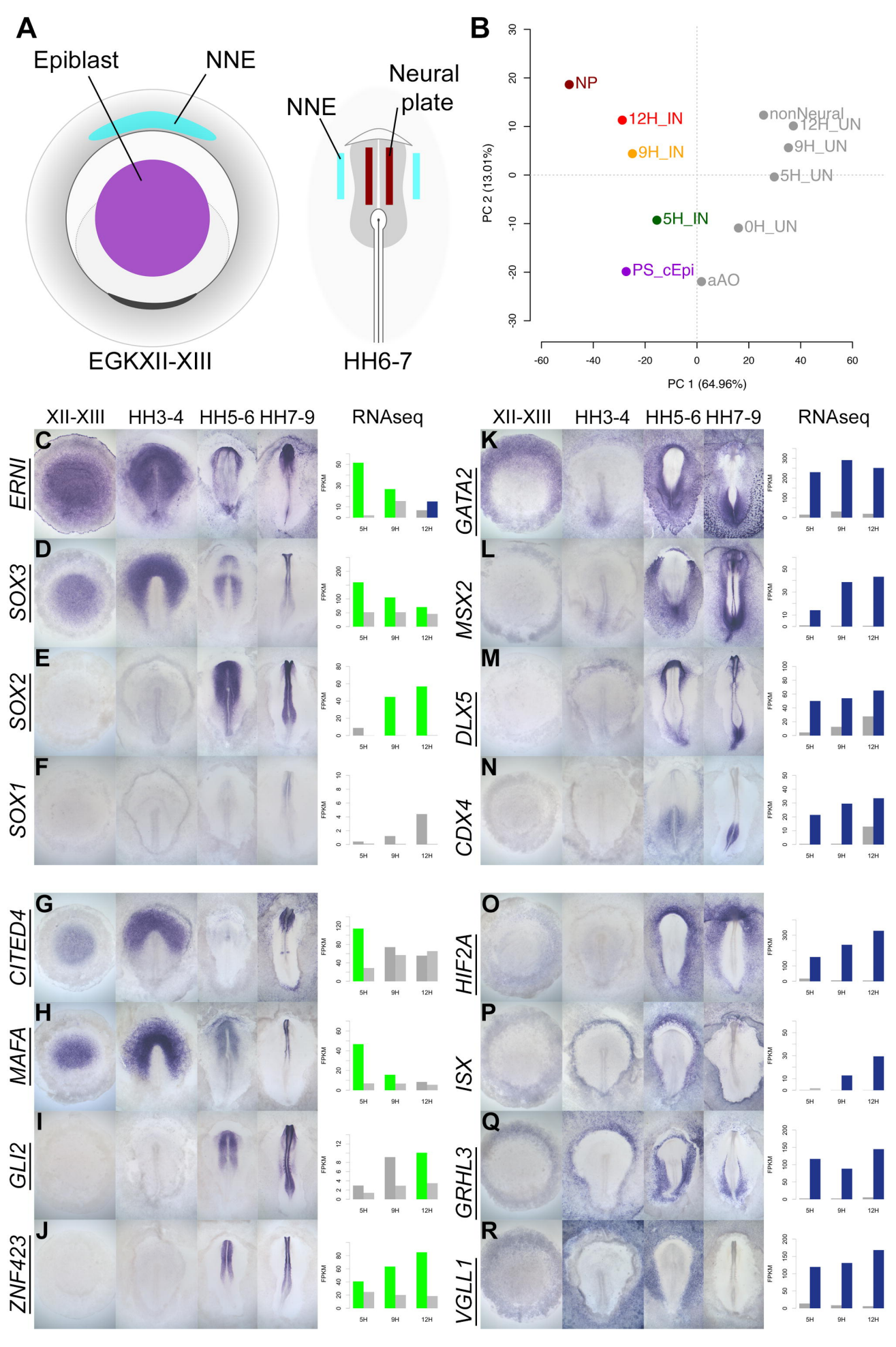

A
HN
HH3-4
HH3-4
15h
SOX2
OTX2
HOXB1
XII-XIII
HH3-4
HH5-6
HH7-9
HN graft
OTX2
SOX2
SOX1
B
C
D
E
F
3h
7/7
G
H
I
J
K
9h
8/8
L
M
N
O
P
12h
3/6
Q
0h UN
HN
HH4-
HH4-
5/9/12h
IN
UN
↑4130 ↓4543
R
S
T
-log10 p-value
log2 FC
5h
9h
12h
U
255
24
203
Transcriptional Regulators
V
5h
9h
12h
35
19
49
53
13
44
66
UP
W
5h
9h
12h
36
6
23
67
8
36
51
DOWN

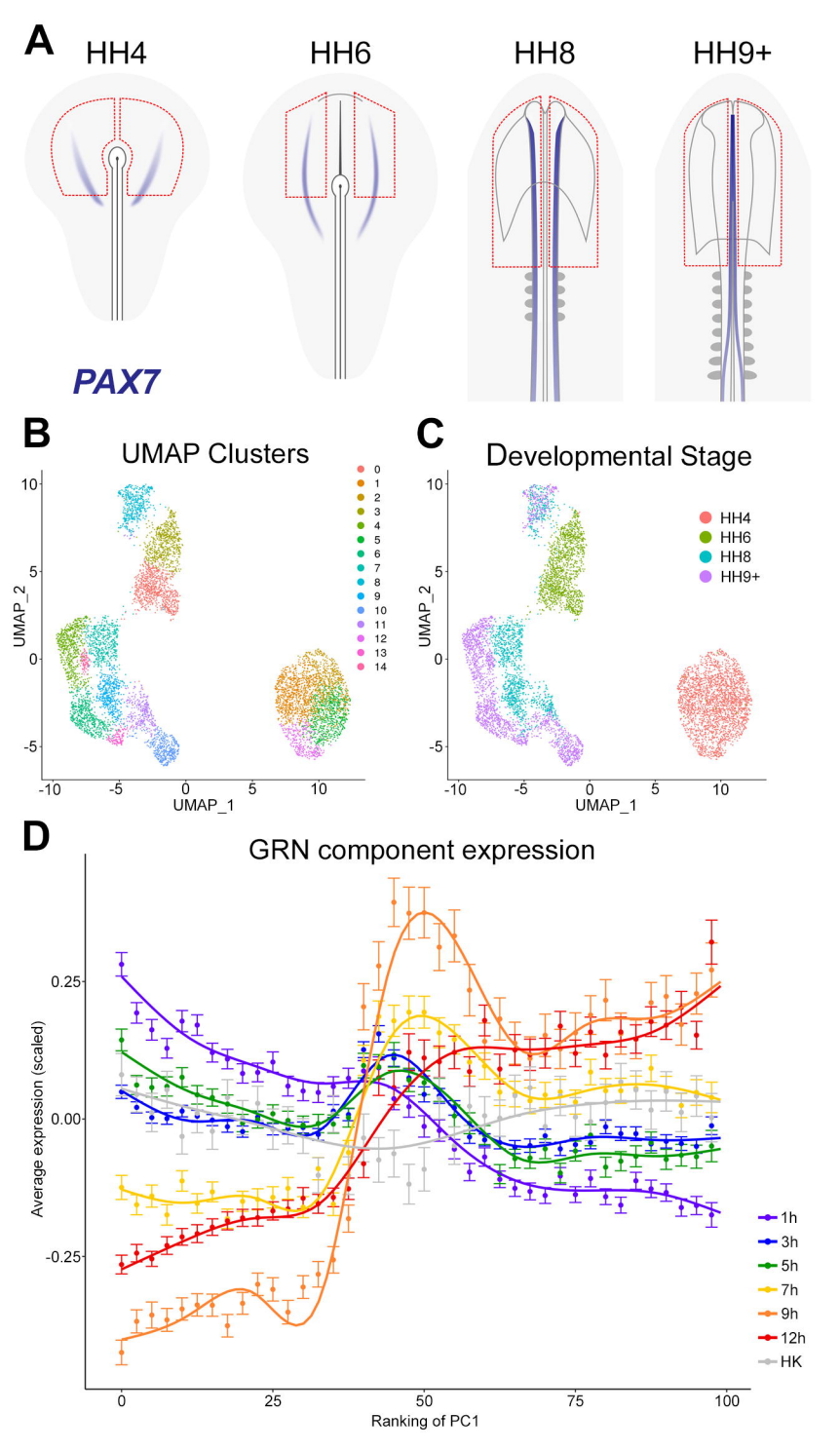

A
0h UN
HN
HH4-
HH4-
5/9/12h
IN
UN
H3K27ac
H3K27me3
B
5h
Induced
Uninduced
Activation
Repression
Poised
No Change
-2.0 PC 2.0Kb
Activation
Poised
Repression
No Change
C
9h
D
12h
H3K27ac
H3K27me3
IN_ac
UN_ac
IN_me3
UN_me3

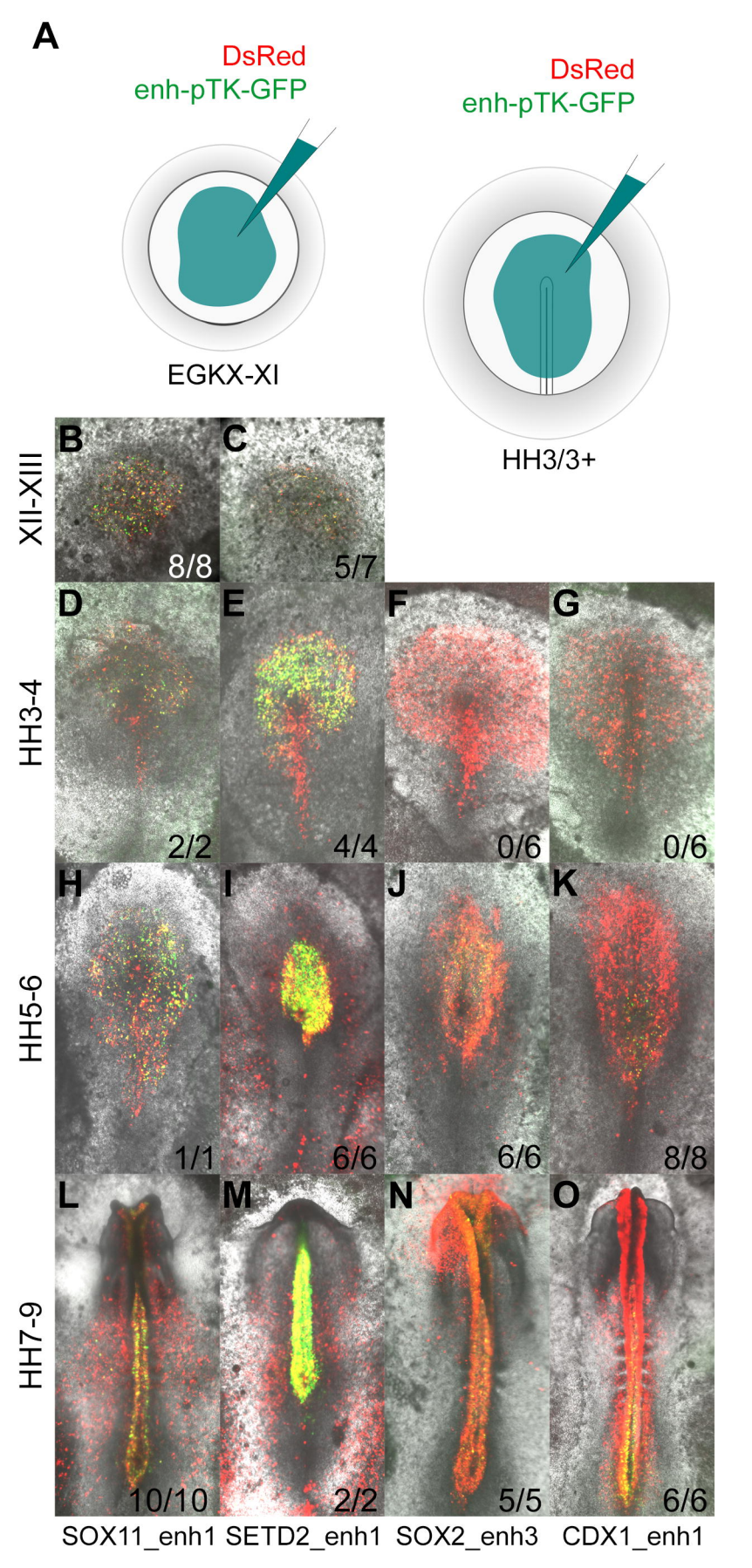

A
ChIPseq
pre-processed data
5/9/12h
CTCF ChIPseq
pre-processed data
Kadota et al., 2017
RNAseq
pre-processed data
0/5/9/12h
DESeq
Identify putative regulatory regions
Identify CTCF sites
Differential gene expression analysis
8673 differentially expressed genes
5/9/12h
Curated TF binding consensus library
Identify NI-specific regulatory regions
Expression data
0/5/9/12h
Select differentially expressed TRs
482 TRs
180 TRs
MEME Suite
FIMO
NI cis-regulatory element info
Merge RNAseq & NanoString data
NanoString
213 TRs
TF binding sites
DREiVe
Expression data
0/1/3/5/7/9/12h
Expression data
1/3/5/7/9/12h
180 TRs
ATACseq
pre-processed data
0/5/9/12h
Computationally predicted regulatory regions
175 TRs
BioTapestry
input files
Generate UCSC browser input files
Build a GRN
*.bed & *.bw files
Data files
In-house scripts
Standard software


B
GRN: 175 components; 5,614 interactions


A
Scale
chr13:
20 kb
galGal5
10,030,000
10,040,000
10,050,000
10,060,000
10,070,000
10,080,000
10,090,000
CTCF
CTCF DRX082681 Gg-embryo ChIPseq
node graft 5H gene expression
node graft 9H gene expression
node graft 12H gene expression
RNAseq
Ensembl Gene Predictions - archive 92 - apr2018
SFXN1
DRD1
BRD8
NHP2
N4BP3
KIF20A
node graft 5H TF binding sites
node graft 5H cis-regulatory sites
node graft 9H TF binding sites
node graft 9H cis-regulatory sites
node graft 12H TF binding sites
node graft 12H cis-regulatory sites
5H_TF_binding
5H_regulatory_sites
9H_TF_binding
9H_regulatory_sites
12H_TF_binding
12H_regulatory_sites
ChIPseq
GRN sites
-
S1
S2
S3
S4
S5
S6
S7
Active
Repressed
No change
Binding sites:
TF expressed
TF not expressed
B
SOX4
OTX2
SNAI1
MGA
ETV5
MEIS2
ZIC3
SOX3
EOMES
HIF1A
MYCN
HEY1
TCF12
SALL1
GBX2
HES1
ZIC2
FOXA2
LHX5
SP4
HESX1
ZNF821
LMX1B
SIX3
HOXB1
5h
BRD8
BRD8_site1
BRD8_site2
BRD8_site4
BRD8_site5
BRD8_site6
BRD8_site7
C
9h
D
12h


A
Epiblast
NNE
EGKXII-XIII
NNE
Neural plate
HH6-7
B
NP
12H_IN
9H_IN
5H_IN
PS_cEpi
nonNeural
12H_UN
9H_UN
5H_UN
0H_UN
aAO
PC 2 (13.01%)
PC 1 (64.96%)
XII-XIII
HH3-4
HH5-6
HH7-9
RNAseq
C ERNI
D SOX3
E SOX2
F SOX1
G CITED4
H MAFA
I GLI2
J ZNF423
K GATA2
L MSX2
M DLX5
N CDX4
O HIF2A
P ISX
Q GRHL3
R VGLL1
FPKM
5H
9H
12H


A
HH4
HH6
HH8
HH9+
PAX7
B
UMAP Clusters
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
UMAP_2
UMAP_1
C
Developmental Stage
HH4
HH6
HH8
HH9+
UMAP_2
UMAP_1
D
GRN component expression
Average expression (scaled)
Ranking of PC1
1h
3h
5h
7h
9h
12h
HK


A
DsRed
enh-pTK-GFP
EGKX-XI
DsRed
enh-pTK-GFP
HH3/3+
XII-XIII
HH3-4
HH5-6
HH7-9
B
8/8
C
5/7
D
2/2
E
4/4
F
0/6
G
0/6
H
1/1
I
6/6
J
6/6
K
8/8
L
10/10
M
2/2
N
5/5
O
6/6
SOX11_enh1
SETD2_enh1
SOX2_enh3
CDX1_enh1