A gradient border model for cell fate decisions at the neural plate border

Alexandre Thiery[1], Ailin Leticia Buzzi[1$], Eva Hamrud[1$], Chris Cheshire[2], Nicholas Luscombe[2#], James Briscoe[2], Andrea Streit[1*]

1 Centre for Craniofacial and Regenerative Biology, Faculty of Dentistry, Oral and Craniofacial Sciences, King's College London, London, UK.

2 Bioinformatics and Computational Biology Laboratory, The Francis Crick Institute, 1 Midland Road, London, UK.

Current address: Genomics and Regulatory Systems Unit, Okinawa Institute of Science and Technology Graduate University, Japan.

$These authors contributed equally to this work.

*Corresponding author: andrea.streit@kcl.ac.uk

Abstract

The vertebrate 'neural plate border' is a transient territory at the edge of the neural plate containing precursors for all ectodermal derivatives: the neural plate; neural crest; placodes; and epidermis. Elegant functional experiments in a range of vertebrate models have provided an in-depth understanding of gene regulatory interactions within the ectoderm. However, these experiments conducted at tissue level have raised seemingly contradictory models for fate allocation of individual cells. Here, we carry out single cell RNA sequencing of chick ectoderm from primitive streak to neurulation stage, to explore cell state diversity and heterogeneity. We identify and characterise the dynamics of gene modules containing key factors known to regulate ectodermal cell fates, allowing us to model the order in which these fates are specified. Furthermore, we find that genes previously classified as neural plate border specifiers typically exhibit dynamic expression patterns and are biased towards either placodal or neural crest fates. Instead, given its transient and heterogenous characteristics, the neural plate border should be defined based on the co-expression of alternative placodal and neural crest gene modules. Finally, we propose a 'gradient border' model for cell fate choice at the neural plate border, with the probability of cell fate allocation closely tied to the spatiotemporal positioning of cells.

Introduction

The vertebrate nervous system is arguably the most complex organ of our body. During development, it arises from only three progenitor populations: the neural plate generates the entire central nervous system, while neural crest and sensory placodes jointly form the peripheral nervous system. Placode derivatives are confined to the head contributing to sense organs and cranial ganglia, while neural crest cells form along the entire body axis to generate various cell types including neurons and glial as well as parts of the craniofacial skeleton, cartilage and teeth (for review: Baker & Bronner-Fraser, 1997; Baker & Bronner-Fraser, 2001; Grocott et al., 2012; Patthey et al., 2014; Pla & Monsoro-Burq, 2018; Prasad et al., 2019; Simoes-Costa & Bronner, 2015).

The neural plate becomes morphologically distinct shortly after gastrulation. However, neural, neural crest and placode precursors continue to be intermingled at its edge, in a territory termed the neural plate border (NPB) that is characterised by the co-expression of neural and non-neural markers as well as so-called NPB specifiers (Pla & Monsoro-Burq, 2018; Roellig et al., 2017; Streit, 2002; Thiery et al., 2020). As the neural plate invaginates, placode precursors remain in the surface ectoderm, while neural crest cells are largely incorporated into neural folds from where they delaminate and migrate to their final destination. Although the induction of these three fates has been studied for decades, some of the most fundamental questions remain unresolved. For example, we do not know when and how these lineages segregate, whether they do so in a specific order and how individual cells make these decisions.

Modulation of different signalling pathways plays a crucial role in controlling the fate of all ectodermal cells. FGF signalling is required for the specification of neural, neural crest and placode precursors and for positioning the NPB (Ahrens & Schlosser, 2005; Litsiou et al., 2005; Londin et al., 2005; Monsoro-Burq et al., 2003; Streit et al., 2000; Streit & Stern, 1999; Stuhlmiller & García-Castro, 2012; Wilson et al., 2000; Yardley & Garcia-Castro, 2012). While BMP signalling must be inhibited for neural development (Hemmati-Brivanlou et al., 1994; Linker & Stern, 2004; Londin et al., 2005; Sasai et al., 1995), it is necessary for placode precursor formation at gastrulation, but must be switched off thereafter (Ahrens & Schlosser, 2005; Kwon et al., 2010; Litsiou et al., 2005). Likewise, neural crest cells need BMP activity early, but unlike placodes continue to do so as they are specified (Barth et al., 1999; Kwon et al., 2010; Liem et al., 1995; Marchant et al., 1998; Steventon et al., 2009). Finally, Wnt activity is required for neural crest cell induction (Garcia-Castro et al., 2002; Steventon et al., 2009) but must be inhibited for neural and placode precursors to form (Heeg-Truesdell & LaBonne,

2006; Litsiou et al., 2005; Patthey et al., 2008). These signals activate different transcription factor networks in a temporally and spatially controlled manner that in turn impart cell identity and have therefore been termed fate 'specifiers'. For example, *Msx1* and *Pax3* have been implicated as NPB specifiers (Milet et al., 2013; Monsoro-Burq, 2015; Plouhinec et al., 2014), while *Pax7* (Basch et al., 2006) and *Six1* (Brugmann et al., 2004; Chen et al., 2009; Christophorou et al., 2009; Ozaki et al., 2004) are involved in neural crest and placode precursor specification, respectively. These factors appear to act in feed forward loops, upregulating their own expression while simultaneously repressing factors that specify alternative fates (for review: Grocott et al., 2012; Thiery et al., 2020).

Based on these studies combined with tissue transplantation experiments, two apparently contradictory models for the segregation of neural, neural crest and placodal fates have emerged. The 'binary competence model' suggests that an initial step in lineage segregation entails a subdivision of the ectoderm into neural/neural crest and placode/epidermal territories (Pieper et al., 2012; Schlosser, 2014). This implies a closer relationship of neural crest cells with the central nervous system and of placodes with the epidermis. In contrast, the 'NPB model' suggests that cells at the border of the neural plate have mixed identity and retain the ability to generate all ectodermal derivatives until after neurulation begins (Roellig et al., 2017; Schlosser, 2014; Thiery et al., 2020). Studies in *Xenopus* have taken this idea a step further, revealing that unlike other ectodermal cells, neural crest or NPB cells retain pluripotency markers and that this is key for their ability to contribute to multiple fates (Buitrago-Delgado et al., 2015; Prasad et al., 2012). In support of this idea, the NPB region is initially characterised by co-expression of neural and non-neural markers (for review: Grocott et al., 2012; Pla & Monsoro-Burq, 2018; Thiery et al., 2020). Fate specifier genes are activated within this territory in partially overlapping patterns to resolve into distinct domains as neurulation begins. While both models are attractive and provide working hypotheses, the underlying experiments were analysed at cell population, rather than at single cell level. They therefore do not address how individual NBP cells make decisions, nor whether they are multipotent or a population of mixed progenitors that are already biased towards their later identity. Recent findings have proposed that prior to fate commitment cells show increasing levels of heterogeneity in gene expression and do not exhibit a definitive transcriptional signature (Soldatov et al., 2019). It is therefore possible that NBP cells represent such undetermined progenitor cells.

Recent studies in chick have begun to explore these questions. Single cell analysis at the NPB reveals considerable heterogeneity of gene expression: factors previously thought to be fate-specific are co-expressed in individual cells even as late as neural tube stages (Roellig et al., 2017; Williams et al., 2022). Furthermore, lineage tracing using a 'neural-specific' Sox2

enhancer shows that Sox2$^+$ cells can contribute not only to neural tissue but also to neural crest and epidermis (Roellig et al., 2017). Together, these findings suggest that cell fate allocation at the NPB occurs over a much longer developmental period than previously thought and that NPB cells retain plasticity until early neurulation. However, although some factors have been identified, the underlying molecular mechanisms of how individual cells make decisions at the NBP remain poorly understood.

Here, we explored the transcriptional changes as cells transit from the neural plate border at early primitive streak to definitive neural, neural crest and placodal fates at single cell level. In total we present single cell RNAseq (scRNA) data for 17,992 high quality cells. Rather than using 2-3 molecular markers to assign cell identity, we established a binary knowledge matrix using previously published *in situ* hybridisation expression data. This allows us to classify cells at each developmental stage, despite the high levels of heterogeneity observed. We find that prior to head fold stages cells do not fall into transcriptional distinct groups but are characterised by graded and partially overlapping gene expression. Their organisation largely reflects their position along the anterior-posterior and medial-lateral axis in the embryo. As expected, cell diversity increases as development proceeds, with a division of neural and non-neural fates becoming apparent at the first somite stage together with the emergence of neural crest cells. By the 8-somite stage, neural, neural crest and placodal cells are largely segregated, although even at this late stage we find cells with heterogenous expression profiles.

Our analysis reveals that NPB cells can already be identified at primitive streak stages by the overlapping expression of neural and non-neural genes, suggesting that this cell state is established before NPB specifiers are activated. Using a gene module approach, we identify groups of genes with dynamic changes in our dataset that are characteristic for neural, neural crest and placodal fates. Comparing co-expression of whole gene modules (rather than individual transcripts) reveals that NPB cells co-express gene modules which later become restricted to placodal or neural crest states. This suggests that NPB cells are in an undecided transcriptional state and may have the potential to give rise to both (or more) fates. Interestingly, although we did not find NPB cells as defined by classical NPB markers at ss4, we did identify cells that co-express competing placodal and neural crest programmes. These cells connect segregating cell populations, resemblant of previously characterised bridge cells in the differentiating murine neural crest which exhibit high heterogeneity prior to segregation (Soldatov et al., 2019). Here, we propose that ‘bridge cells’ between placodal and neural crest populations continue to be plastic, maintaining the ability to contribute to different fates. Overall, our analysis has led us to revisit the definition of the NPB and propose that NPB cells

may be better defined by the co-expression of competing gene modules rather than a particular gene signature.

Results

Classification of ectodermal cells using a binary knowledge matrix of known marker genes

Neural, neural crest and placode progenitors arise from the embryonic ectoderm. However, how individual cells diversify and the sequence and timing of these events remain to be elucidated. To capture their segregation, we carried out 10x Genomics single-cell mRNA sequencing (scRNAseq) on cells taken from a broad region of the chick epiblast at six developmental stages: HH4^{-}/4 (primitive streak), HH5 (head process), HH6 (head fold), HH7 (1 somite), somite stage (ss) 4 and ss8 (Figure 1A). These stages were chosen because they encompass the initial formation of the neural plate to the beginning of neurulation when these fate decisions are thought to take place. After integration across batches, stringent quality control and removal of contaminating cell populations (see methods and Figure 1–figure supplement 1A-I), we captured a total of 17,992 high quality cells with approximately 2500 – 4000 cells per stage and an average of 3924 genes per cell. Dimensionality reduction and UMAP embedding showed that cells were organised primarily by stage (Figure 2E). Later stages displayed higher cell diversity as compared to earlier stages where most cells were transcriptionally similar (Figure 1B-E and Figure 2A-D).

Typically, scRNAseq cell states are classified using a small number of definitive marker genes which exhibit mutually exclusive expression patterns. However, this approach is difficult to apply to the NPB for several reasons. Firstly, the early NPB cannot be classified using definitive markers as it emerges in the region of overlap between neural and non-neural genes. Secondly, the NPB is a transitory state through which cells are thought to pass prior to acquisition of any of the four ectodermal fates (neural, neural crest, placodal and epidermal). Finally, cells at the NPB appear to be highly heterogeneous, with precursors for the different ectodermal fates found intermingled, even at late stages of neurulation (Bhattacharyya et al., 2004; Roellig et al., 2017; Streit, 2002; Williams et al., 2022). To overcome these challenges, we categorised cells using a binary knowledge matrix based on expression data obtained from the literature (Supplementary File 1). We examined available *in situ* hybridisation data of 76 genes with known regionalised expression patterns and binarized their expression within each

defined cell state. This allowed us to define a total of 24 cell states spanning primitive streak to neural tube stages.

We first unbiasedly clustered cells from each developmental stage at a high resolution to ensure capturing the full diversity of cell states (Figure 1–figure supplement 1J). Clusters were then classified based on the similarity between their gene expression and the knowledge matrix (see methods). Next, cells from all stages were clustered together and cell state labels transferred to the entire data set (Figure 2F). This classification approach successfully identified 22 distinct cell states within the ectoderm. As expected, cell diversity increases as development proceeds. Early NPB and early neural cell states can already be identified at primitive streak stages, while placode progenitor and neural crest states emerge at head process and ss4, respectively (Figure 1B and Figure 2A-B).

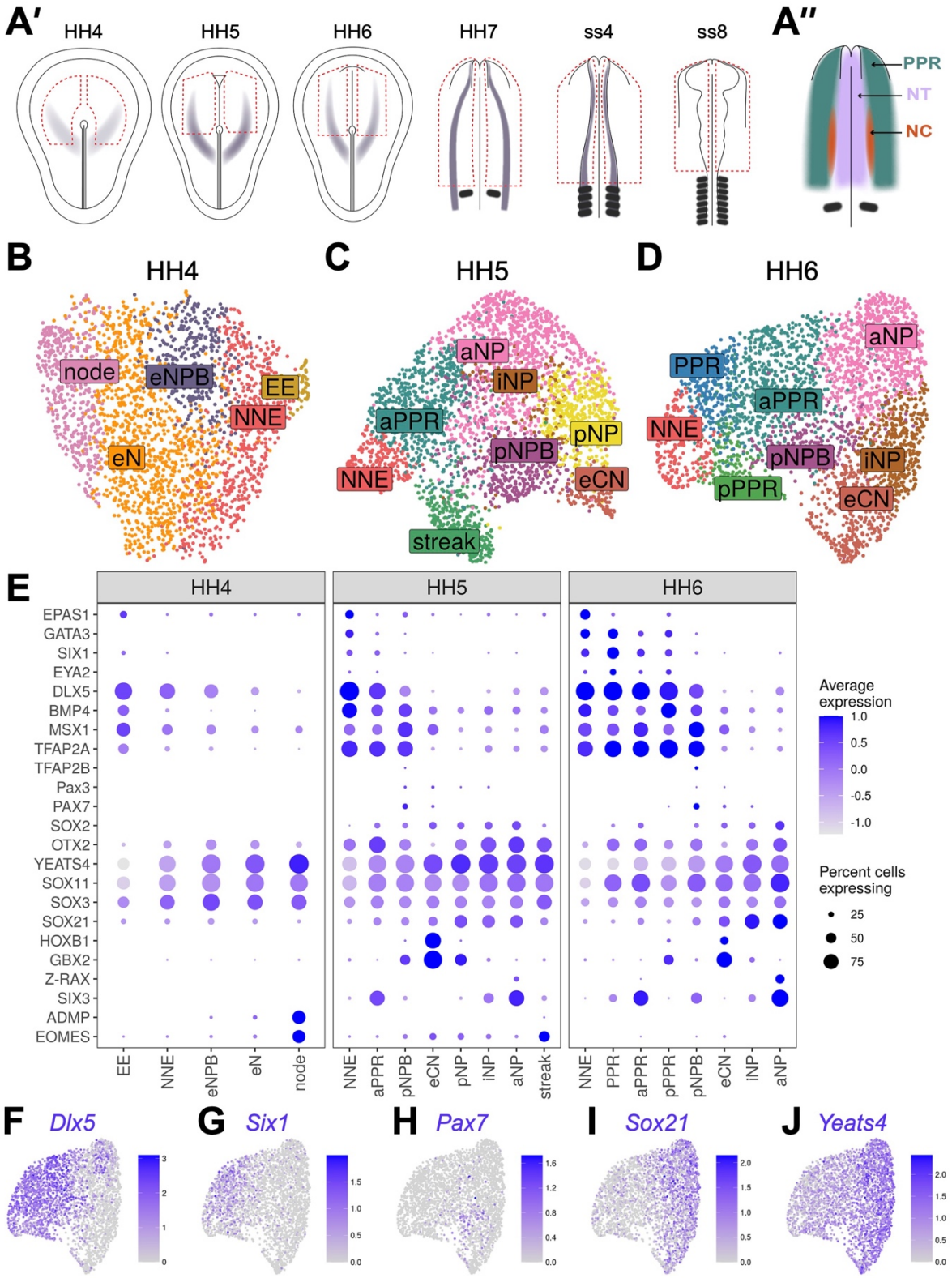


Figure 1. Cells at stages HH4 – HH6 reflect the anterior-posterior and medio-lateral axes in the embyro. (A') Dorsal view schematics of chick embryo at stages HH4-ss8 depicting the ectodermal tissue region dissected for 10x scRNAseq (red dotted line). Purple shading illustrates previously characterized region of *Pax7* expression (Basch et al. 2006). **(A'')** Schematic of a 1-somite stage (HH7) chick embryo illustrating the pre-placodal region (PPR),

neural crest (neural crest) and neural tube (NT). **(B-D)** UMAP plots for cells collected at HH4^{-}/4 (primitive streak), HH5 (head process), HH6 (head fold) stages. Cell clusters are coloured and labelled based on their semi-unbiased cell state classifications (binary knowledge matrix available in Supplementary File 1). **(E)** Dot plots displaying the average expression of key marker genes across cell states at HH4, HH5 and HH6. The size of the dots represents the number of cells expressing a given gene within each cell population. **(F)** Feature plots revealing the mediolateral segregation of key marker genes at HH6.

Cell states emerging from primitive streak to head fold stages

To explore the changes in cell diversity within the embryonic ectoderm over time, we compared cell states present at different stages and identified the time point when each emerged. We visualised cell transcriptomes in low dimensionality space using UMAPs, in turn allowing us to infer relationships between cell states and explore the distribution of gene expression across the dataset. We first examined stages HH4-HH6 to characterise events as the neural plate is established and placodal and neural crest cells are thought to be specified.

An early NPB state is identified at primitive streak stages prior to the expression of NPB specifiers

By primitive streak stage, mediolateral (M-L) gradients of molecular markers demarcate pre-neural and non-neural territories within the chick epiblast (Pera et al., 1999; Rex et al., 1997; Sheng & Stern, 1999; Streit et al., 2000; Streit et al., 1998). At HH4, unbiased clustering clearly organises cells in a pattern reminiscent of this M-L axis (Figure 1B; from left to right) as evidenced by the expression of node (*ADMP*), pre-neural (*Yeats4, Sox3, Sox11*) and non-neural (*Dlx5*, *Bmp4, Msx1, Tfap2a*) genes (Figure 1B, E and Figure 1–figure supplement 2A). Cell states can be broadly characterised based on combinatorial gene expression (Figure 1E) and we identify five cell states: node, early neural (eN), early neural plate border (eNPB), non-neural ectoderm (NNE) and extraembryonic ectoderm (EE) (Figure 1B). While some genes are shared between many cells, early neural plate cells are defined by high levels of pre-neural transcripts while the non-neural and extraembryonic ectoderm are characterised by increased non-neural gene expression. Notably, we find that an early NPB state can be identified by the co-expression of both neural and non-neural genes. Thus, NPB cells are already apparent prior to the differential upregulation of genes previously characterised as NPB specifiers (*Msx1, Pax7, Tfap2a, Bmp4, Dlx5, Pax3*) (Basch et al., 2006; Ezin et al., 2009; Hong & Saint-Jeannet, 2007; McLarren et al., 2003; Rothstein & Simoes-Costa, 2020; Streit & Stern, 1999) (for review: Simoes-Costa & Bronner, 2015; Thiery et al., 2020).

The first emerging NPB and PPR cells have regional identity

205 At HH5, cell positions in the UMAP continue to reflect organisation along the M-L axis in the
206 embryo (Figure 1C), but new transcriptional states emerge. Whilst non-neural ectoderm cells
207 display a similar expression profile at HH5 compared to HH4, neural plate cells differ from
208 early neural cells in their expression of definitive neural (e.g., *Sox2*) and regional markers
209 (Figure 1E; Figure 1–figure supplement 2B). Like at HH4, many cells continue to co-express
210 neural and non-neural transcripts (Figure 1–figure supplement 2B). Within the latter cell group,
211 we observe the upregulation of PPR genes (*Six1*, *Eya2*; Figure 1C, E: aPPR, anterior PPR)
212 and for the first time NPB specifiers (*Msx1*, *Pax3, Pax7*; Figure 1C, E: pNPB, posterior NPB).
213 These observations correlate well with gene expression in the embryo. *Msx1* is initially
214 expressed in the extraembryonic and non-neural ectoderm, and together with *Pax7* begins to
215 be upregulated in the NPB thereafter (Basch et al., 2006; Streit & Stern, 1999), while PPR
216 markers are first expressed at HH5. Notably, the first NPB cells to emerge have posterior
217 character, while placode precursors have anterior identity based on their expression of *Gbx2*
218 and *Six3*, respectively (Figure 1E; see also below).
219
220 *Anteroposterior organisation of cells at head process and head fold stages*
221 Cells with distinct anteroposterior (A-P) character are first observed at HH5 through the
222 restricted expression of various homeobox transcription factors. *Rax* and *HoxB1* are highly
223 confined to distinct cell groups, whilst in contrast *Six3* and *Gbx2* exhibit broader negatively
224 correlated expression gradients across cell clusters with some degree of overlap (Figure 1E
225 and Figure 1–figure supplement 2B). The organisation of cells within the UMAP reflects the A-
226 P expression gradients of these genes in the embryo (Chapman et al., 2002; Hidalgo-Sanchez
227 et al., 2005; Ohuchi et al., 1999; Paxton et al., 2010). The overlap of A-P factors was used to
228 distinguish different NP, PPR and NPB regions during unbiased cell state classification
229 (Supplementary File 1).
230
231 Within the NP we can distinguish intermediate neural plate (iNP) states from the anterior and
232 posterior neural plate (aNP, pNP) and from early caudal neural (eCN) states. However, within
233 the PPR and NPB we identify only anterior (aPPR; *Six3*$^{+}$) and posterior (pNPB; *Gbx2*+) cell
234 states, respectively. While in UMAPs aPPR cells are located between the NNE and anterior
235 NP, the pNBP cluster resides between the iNP and pPPR clusters (Figure 1C, E Figure 1–
236 figure supplement 2C), reflective of both its A-P and M-L position *in vivo*.
237
238 Between HH5 and HH6, posterior *Gbx2*$^{+}$ PPR (pPPR) cell states emerge (Figure 1D), as do
239 PPR cells that appear to lack regional identity. At this stage, *Otx2* and *Gbx2* expression
240 domains in the embryo abut (Steventon et al., 2012), suggesting that cells with generic PPR
241 character might be intermingled with those already biased towards axial identity and either

242 retain the potential to contribute to multiple axial levels or will later become intermediate PPR
243 cells.
244
245 Together, our data show that the emergence of cells with divergent transcriptional profiles
246 resembles patterning of the ectoderm *in vivo*. However, unbiased cell clustering forces cells
247 into discrete clusters which may not fully reflect heterogeneity in gene expression and cell
248 diversity. Indeed, our single cell data reveal the heterogenous nature of gene expression:
249 markers are broadly expressed and are not confined by cluster boundaries or by neighbouring
250 expression domains (Figure 1F-J, Figure 1–figure supplement 2C). Despite this heterogeneity,
251 our cell classification approach successfully identifies increased ectodermal cell diversity over
252 time, with neural, posterior NPB and anterior PPR cells emerging when the definitive neural
253 plate forms. Importantly, the relative positions of these cell states in the UMAPs reflect their
254 location in the embryo.
255
256 **Increasing cell diversity from 1-somite to 8-somite stages**
257
258 At HH7, the neural folds begin to elevate indicating the start of neurulation, while during early
259 somitogenesis neural crest cells start to migrate and placode precursors begin to diversify.
260 We therefore sought to characterise the transcriptional changes at single cell level as neural,
261 neural crest and placodal fates emerge.
262
263 *Emergence of neural crest cells and segregation of neural/non-neural fates*
264 Following clustering, we identified two major superclusters at HH7, one containing neural tube
265 and neural crest and the other placodal and future epidermal clusters (Figure 2A). These
266 superclusters appear to be largely separate, both by their UMAP embedding and the restricted
267 expression of definitive fate markers (Figure 2A, D). Whereas at earlier stages many cells co-
268 express non-neural and neural markers (Figure 1E, F, J), at HH7 cells expressing non-
269 neural/placodal (*Dlx5*, *Six1*) and neural transcripts (*Sox21*, *Sox2*) are largely confined to their
270 respective supercluster (Figure 2–figure supplement 1A). It is at this stage that anterior NPB
271 cells can first be identified transcriptionally, and together with posterior NPB cells they connect
272 the two superclusters (Figure 2A; aNPB, pNPB).
273
274 Adjacent to the NPB clusters, we observe a neural crest cell cluster expressing early definitive
275 neural crest markers including *Tfap2b*, *Snai2* and *Axud1* (*Csrnp1*) alongside NPB markers
276 (*Msx1*, *Pax7* and *Tfap2a*) (Figure 2D; NC and Figure 2–figure supplement 1A). Thus, neural
277 crest cells first emerge as a distinct cell population at HH7. Although many genes are shared
278 between the NPB and neural crest cell clusters, NPB cells also express PPR and non-neural

markers (*Six1*, *Eya2* and *Dlx5*) which are in turn downregulated in neural crest cells (Figure 2D and Figure 2–figure supplement 1A). Furthermore, we find that *Msx1* and *Pax7* are never strongly expressed within placodal cell clusters. It is therefore possible that these NPB specifiers may not represent NPB markers, but rather an early bias towards neural crest fate.

'Bridge cells' identified as putative undetermined cells undergoing fate decisions during neurulation

At ss4 and ss8, NPB cells can no longer be identified using the above criteria. Instead, we observe an increasing diversity of cells classified as placodal, neural crest or neural (Figure 2B, C). Within the neural cell population, forebrain, midbrain, and hindbrain (FB, MB, HB) clusters are clearly distinguishable. *Six3* and *Gbx2* expression define cells with fore- and hindbrain identity, respectively, whilst the upregulation of the dorsal midbrain marker *Wnt4 (Hollyday et al., 1995)* and mid/hindbrain boundary marker *Pax2* (Hidalgo-Sanchez et al., 1999) defines the midbrain-like cells (Figure 2D and Figure 2–figure supplement 1B). As development progresses, neural, neural crest and placode cells become increasingly distinct, with neural crest cells beginning to express a suite of definitive neural crest markers, including *Sox10* in the delaminating neural crest (dNC) cluster (Figure 2D and Figure 2–figure supplement 1B, C).

Despite the apparent segregation of neural, neural crest and placodal fates during ss4 and ss8, connecting cells bridge PPR clusters with both neural crest and forebrain clusters, as well as neural crest with midbrain cell clusters in a four-way pattern reflecting their A-P and M-L segregation *in vivo* (Figure 2B, C). To predict the directionality of transcriptional change within these bridge cells, and hence their potential developmental trajectory, we performed RNA velocity analysis (Bergen et al., 2020; La Manno et al., 2018). This analysis predicts that bridge cells move away from cluster boundaries, towards definitive cell states (Figure 2G, H). This observation implies that bridge cells do not have a definitive transcriptional signature and may retain plasticity to contribute to different fates, despite the absence of a distinct NPB state from ss4 onwards.

Together our data show that whereas cell states at HH4-HH6 exhibit broad overlapping gene expression, definitive fate markers are upregulated from HH5 onwards and subsequently gene expression is refined at the boundaries between cell clusters. At HH7 and ss4, placodal, neural crest and neural fates begin to segregate. However, even as the neural tube closes, bridge cells continue to straddle cluster boundaries; these cells may be transcriptionally indeterminate and therefore still actively undergoing fate decision processes.

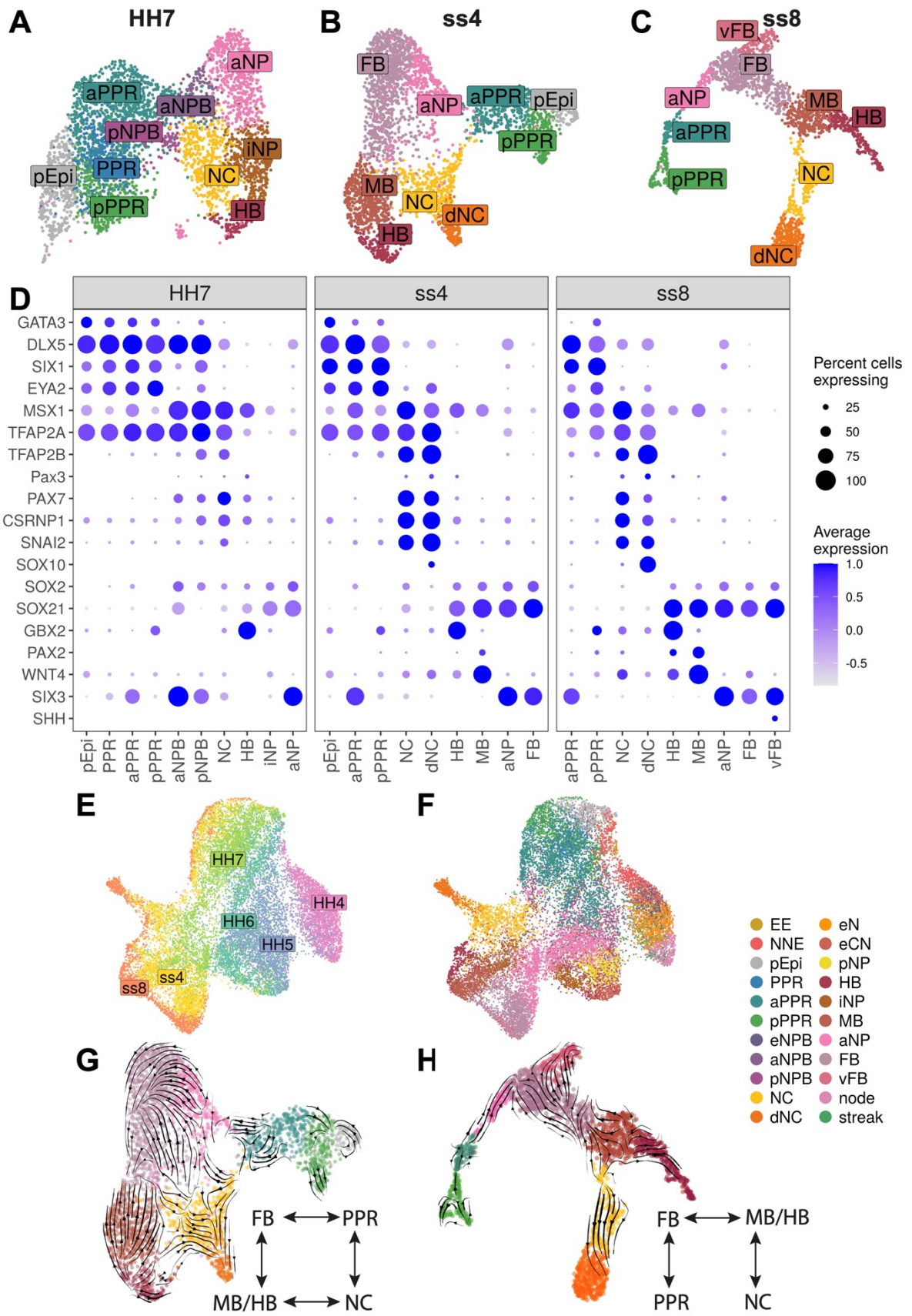


Figure 2. Increased cell diversity and lineage segregation from HH7 to ss8. (A-C) UMAP plots for cells collected at HH7 (1 somite stage), ss4 (4 somite stage) and ss8 (8 somite stage). Cell clusters are coloured and labelled based on their semi-unbiased cell state classifications

(binary knowledge matrix available in Supplementary File 1). **(D)** Dot plots displaying the average expression of key marker genes across cell states at HH7, ss4 and ss8. The size of the dots represents the number of cells expressing a given gene within each cell population. **(E)** UMAP plot for the whole dataset (HH4-ss8 combined) coloured and labelled by stage. **(F)** UMAP plot for the whole dataset coloured by cell state classifications. Cell state classifications were calculated independently for each stage and transferred across for visualisation of the full combined dataset. **(G-H)** UMAP plots for ss4 and ss8 overlaid with RNA velocity vectors depicting the predicted directionality of transcriptional change. Cells are coloured by cell state classification (shown in B and C). Schematics summarise the key fate segregation events taking place at these stages predicted by RNA velocity analysis.

Dynamic changes in gene module expression over time

Many transcriptional regulators and their interactions have been implicated in cell fate decisions as the ectoderm is specified and differentiates into its derivatives. However, in this context studies have largely taken a gene candidate approach. To explore the full transcriptional dynamics of both known and new transcriptional regulators during the specification of neural, neural crest and placodal cells, we modelled the dynamics of whole gene modules rather than individual candidate genes. This allows us to investigate transcriptional dynamics in an unbiased manner whilst simultaneously minimising the loss of potentially important factors.

Defining neural, neural crest and placodal gene modules

To investigate the segregation of neural, neural crest and placode fates, we first sought to identify gene modules that characterise each population. Gene modules are groups of genes that display similar expression profiles across a dataset. We calculated gene modules using the Antler R package (Delile et al., 2019), which iteratively clusters genes hierarchically based on their gene-gene Spearman correlation. This approach reveals 40 modules each showing highly correlated gene expression across all stages (HH4-ss8). To focus on gene modules which may include candidates for fate segregation, we included only those that are differentially expressed between any cell state and the rest of the dataset, and then filtered to remove those with differential expression between sequencing batches. Finally, since our data show that the segregation of neural, neural crest and placodal cell states is almost complete at ss8 (Figure 2C), we kept those gene modules that show differential expression between at least one of the three fates at ss8 (see methods). This resulted in nine modules that were considered further (Figure 3–figure supplement 1A and Figure 3–Source Data 1). Of these, four gene modules (GM21-24) are specifically expressed in the PPR at ss8, one (GM13) in

358 neural crest cells, and the remaining four (GM7, 9, 11, 20) are upregulated primarily in neural
359 clusters (Figure 3–figure supplement 1A).
360
361 When examining the genes within each module, we noticed that GM20 and GM22 include
362 genes with widespread expression or expression in the lateral mesoderm and/or the
363 extraembryonic region (*Epas1, Nkx2-5, Tnnc1, Pgk1*) (Adams et al., 2008; Bell et al., 2004;
364 Liberatore et al., 2000; Ota et al., 2007) (Figure 3–Source Data 1). We therefore removed
365 these modules from subsequent analysis. GM7 contains well characterised anterior markers
366 including *Otx2*, *Six3* and *Pax6*, whilst GM11 contains midbrain/hindbrain markers including
367 *Pax2*, *Nkx6-2* and *Wnt4* (Figure 3–Source Data 1). We find that at ss4 and ss8 these gene
368 modules are indeed upregulated in the forebrain/aPPR and midbrain/hindbrain, respectively
369 (Figure 3–figure supplement 1A). To focus on our analysis on the segregation of neural, neural
370 crest and placodal cells rather than on A-P patterning, A-P restricted modules were removed.
371 The remaining five modules show broadly pan-neural (GM9), pan-placodal (GM23, GM21,
372 GM24) or neural crest cell (GM13) expression at ss4 and ss8 (Figure 3A).
373
374 *Latent time analysis suggests late emergence of neural crest cells*
375 Given the cellular heterogeneity observed at early timepoints, discrete developmental stages
376 do not necessarily accurately represent the developmental progression of individual cells
377 towards a given fate. In contrast, latent time can measure the biological clock of individual
378 cells as they differentiate and thus represents the order of cells along a developmental
379 trajectory. To explore the expression dynamics of the five selected gene modules as neural,
380 neural crest and placodal fates emerge, we first ordered cells across latent time using scVelo
381 for the entire dataset (Figure 3B and Figure 3–figure supplement 1B). To investigate the
382 segregation of all three fates, three separate trajectories need to be distinguished. To do this
383 we calculated the probability of each cell to transit towards one of three terminal states (fate
384 absorption probability) (Figure 3–figure supplement 1C). These probabilities were obtained
385 using CellRank (Lange et al., 2022), which leverages RNA velocity and transcriptomic
386 similarity information to predict transitions between cells. Mapping the fate probabilities of each
387 cell onto UMAP plots of the full dataset (Figure 3C-E, see Figure 3B and Figure 2E-F for
388 reference) shows that cells with low latent time values (i.e., early cells) have low probabilities
389 of becoming neural, neural crest or placodal cells. As cells 'age' their fate probabilities increase
390 as they become more transcriptionally like the terminal state. Unlike placodal and neural,
391 neural crest fate probabilities are only high in 'late' cells (Figure 3B-C), reflecting the fact that
392 defined neural crest cell clusters only appear at HH7.
393

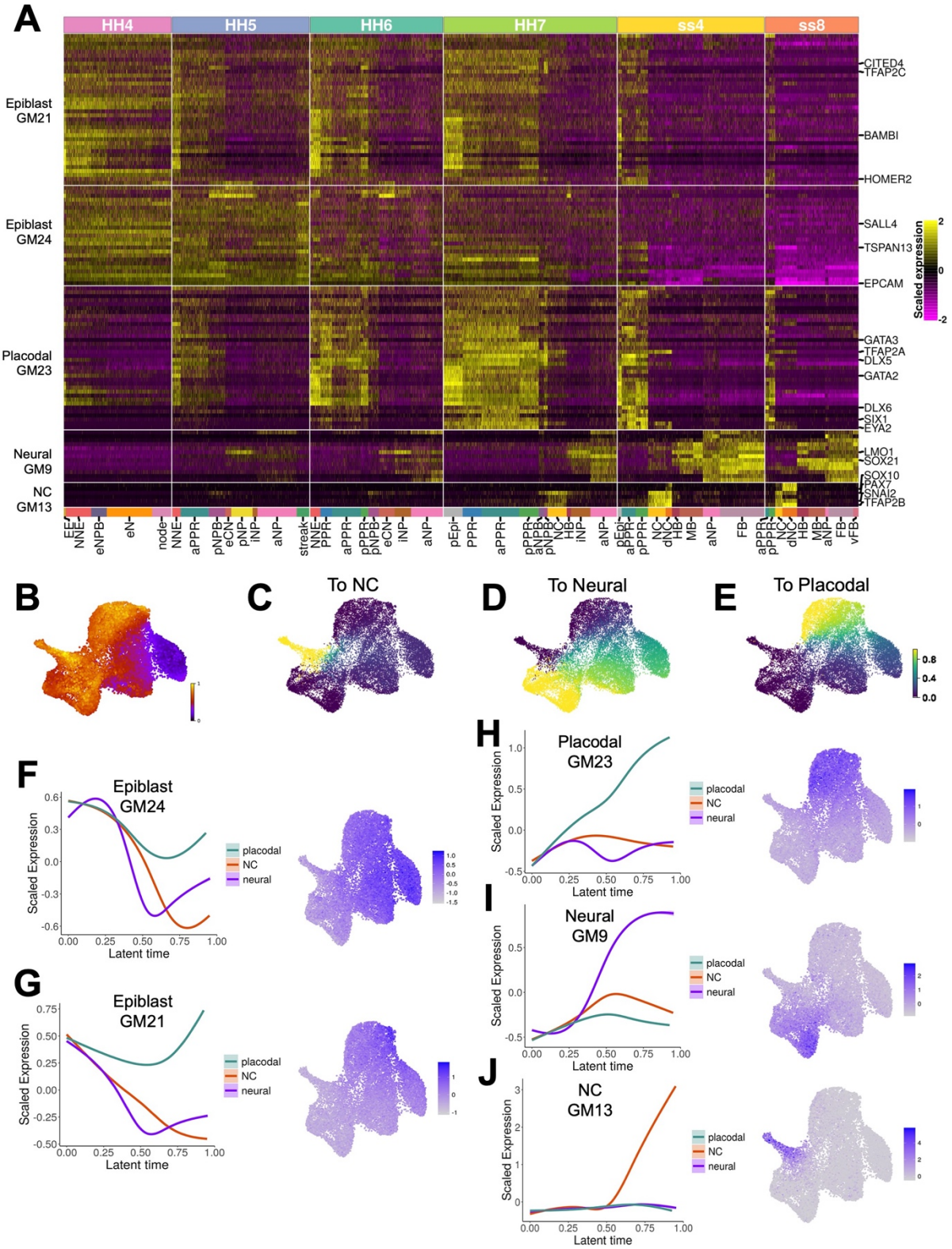


Figure 3. Gene module dynamics reveal key differences between the segregation of the PPR, neural crest and neural fates. (A) Heatmap of gene modules that display fate-specific expression (full list of gene modules available in Figure 3-Source Data 1). Gene modules were first unbiasedly filtered based on their differential expression between cell states (see methods and Figure 3-figure supplement 1A). Gene modules were further manually filtered based on

expression patterns described in the literature (see results section: *Defining neural, neural crest and placodal gene modules).* Key genes of interest are highlighted on the right (for full gene list see Figure 3-Source Data 1). **(B)** UMAP plot of the full dataset (HH4-ss8) showing cell latent time values. **(C-E)** UMAP plots of the full dataset showing the fate absorption probabilities of each cell towards one of the three defined terminal states: neural crest, neural and placodal. **(F-J)** Gene module dynamics plots and feature plots for 5 gene modules: GM24, GM21, GM23, GM13 and GM9. Gene module dynamic plots display generalised additive models (GAMs) of scaled gene module expression across latent time. GAMs are weighted by the fate absorption probability of each cell towards one of the three terminal states (placodal, neural crest and neural). Feature plots show average scaled gene module expression on UMAPs of the full dataset.

Gene module dynamics reveals key differences in the activation of the transcriptional programmes associated to neural, neural crest and PPR specification

Having predicted the developmental trajectory of cells using scVelo and CellRank, we next modelled the expression of the five selected gene modules as a function of latent time with the aim to characterise their dynamics. Gene expression dynamics were modelled using generalised additive models (GAMs). GAMs were chosen because gene expression is highly dynamic and does not change linearly over time. For each gene module we fitted three separate GAMs, one for each fate (Figure 3F-J). The three fates were distinguished using the previously calculated fate absorption probabilities. This modelling approach reveals dynamic changes of gene expression. Three gene modules – GM9, GM13, GM23 – are initially inactive and are upregulated at different points across latent time in cells with neural, neural crest and placodal character, respectively (Figure 3H-J).

GM9 becomes upregulated in all neural clusters and includes neural markers like *Lmo1* and *Sox21* (Figure 3A, I; Figure 3–Source Data 1). Although the expression of different genes within this module correlates broadly, there is some gene-gene variation. For example, *Lmo1* is the first gene to be expressed and exhibits highly dynamic expression over time (Figure 3A). It is first found at HH5 within the eCN, pNP and iNP, before being upregulated across the entire neural plate at HH7 and ss4. At ss8, *Lmo1* is specifically downregulated within the MB. In contrast, *Sox21* is first upregulated at HH5/6 and is consistently expressed throughout all neural cell clusters (Figure 3A).

In contrast, modelling reveals that GM13 is upregulated in cells with neural crest identity (Figure 3J). Indeed, it contains well known neural crest markers *Pax7*, *Snai2*, *Sox10* and *Tfap2b* (Figure 3A, Figure 3–Source Data 1) and is initially broadly activated in the NPB and neural crest clusters at HH7. At ss4 and ss8, this module remains strongly expressed within

neural crest and delaminating neural crest clusters (Figure 3A). Within GM13 individual genes are activated at different time points: *Pax7* is the first gene to be expressed in the pNPB at HH5, while *Sox10* starts to be expressed only from ss4 onwards.

Finally, modelling the expression dynamics of GM23 shows that it is activated early and becomes upregulated in cells with placodal character across latent time (Figure 3H). This prediction is supported by the fact that GM23 is highly expressed in PPR clusters at ss8 and contains *bona fide* placodal genes like *Six1* and *Eya2* (Figure 3A, Figure 3–Source Data 1). At HH4, most GM23 genes are not expressed, but are broadly activated at HH5 in PPR, non-neural and NPB clusters, where they remain active until becoming restricted to PPR cells at ss4 (Figure 3A). Comparing the onset of activation of these three gene modules across latent time reveals a clear temporal order with the placodal module being activated first, followed by the neural and neural crest gene modules.

In contrast to these modules, modelling of GM24 and GM21 shows that they are initially expressed in all three cell populations before being rapidly downregulated within neural crest and neural cells (Figure 3F, G). Indeed, genes in GM24 are broadly expressed across the epiblast at HH4, and then become gradually restricted to PPR, non-neural and NPB clusters at HH6/7, and to PPR cells at ss4/8 (Figure 3A). GM21 displays similar dynamics but is restricted to non-neural clusters somewhat earlier (HH5; Figure 3A). Although some genes within these epiblast modules have been described as non-neural/PPR specific based on *in situ* hybridisation (*Tfap2c*, *Bambi*, *Homer2*, *Tspan13*) (Bell et al., 2004; Hintze et al., 2017; Mehdizadeh et al., 2021; Reichert et al., 2013; Rothstein & Simoes-Costa, 2020), others are found to be expressed within the early neural territory (*Epcam*, *Sall4*, *Cited4*) (Barembaum & Bronner-Fraser, 2007; Bell et al., 2004; Trevers et al., 2021) (Figure 3–Source Data 1).

In summary, modelling the expression dynamics of selected gene modules provides new insight into the sequential activation of fate-specific developmental programmes, as well as proposing key differences in the mechanism by which neural, neural crest and placodal cells are specified. We show, using an unbiased gene module approach, that a placodal gene module is activated first, followed by neural and neural crest modules. While neural and neural crest modules are specifically activated in the relevant cell clusters alone, two modules that are later restricted to the PPR are initially broadly expressed across the epiblast. We therefore propose that the specification of neural and neural crest states may require the downregulation of epiblast modules which later become confined to placodal cells.

477 **Redefining the NPB cell state**

478

479 NPB cells are thought to retain the potential to give rise to neural, neural crest and placodal
480 lineages, and recent single cell analysis suggests that they can do so even at late neurulation
481 stages (Roellig et al., 2017; Williams et al., 2022). So far, our analysis reveals an early NPB
482 cell population identified by the overlapping expression of non-neural and early neural markers
483 at primitive streak stages, albeit being transcriptionally highly heterogeneous. In contrast,
484 *Msx1* and *Pax7* which have previously characterised as NPB specifiers, are only upregulated
485 at the NPB at head process stage, have posterior character and are largely absent from
486 placode precursors (Figure 1E, F, Figure 1–figure supplement 2B, C). At ss4, NPB cells can
487 no longer be identified based on the expression of these specifier genes, yet we find continued
488 presence of bridge cells that connect neural, neural crest and placodal cell clusters. Thus,
489 among the key questions remaining are whether NPB cells have a unique transcriptional
490 profile and whether bridge cells have NPB character with the potential to give rise to multiple
491 fates.

492

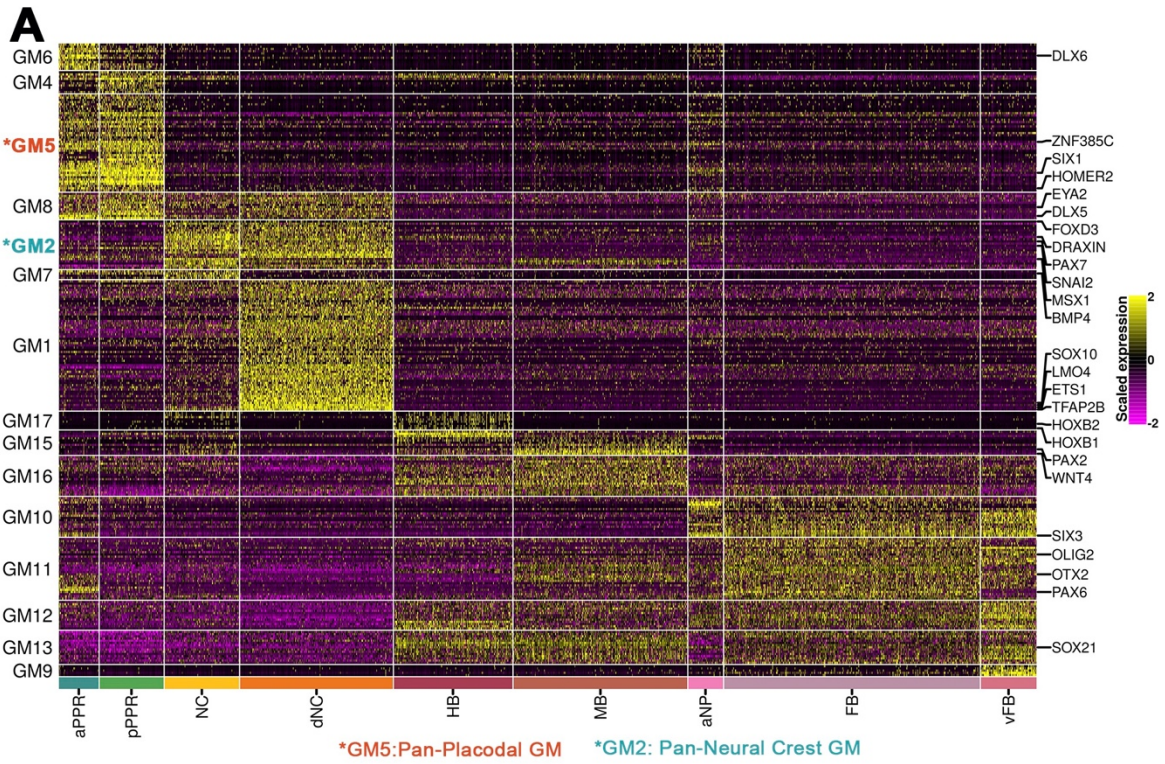
493 To explore this idea, we focused on the decision between neural crest and placode precursors.
494 Neural crest and placode specific transcription factors cross-repress each other while
495 enhancing their own expression (for review: Thiery et al., 2020). It is therefore likely that co-
496 expression of such genes may cells which are still plastic. We therefore examined the co-
497 expression of neural crest and placode specific gene modules in a subset of cells composed
498 of neural crest, placodal and NPB cell states between HH5-ss8 (for workflow explanation see
499 Figure 3–figure supplement 2).

500

501 *Defining neural crest and placode specific gene modules*

502 First, we sought to identify gene modules that are specifically restricted to neural crest and
503 placode cells at late stages. GM analysis in cells collected from ss8 identified 15 modules
504 (Figure 4 and Figure 4–Source Data 1) that are differentially expressed in at least one cell
505 state. GM9-13 and GM16-17 are characteristic for neural cells with different A-P identities.
506 Seven gene modules are absent from neural populations but expressed in placodal cells
507 (GM4-6), neural crest (GM1-2) or both (GM7-8). GM5 contains well characterised placodal
508 markers including *Six1*, *Homer2* and *Znf385c* (Figure 4–Source Data 1) and is expressed in
509 both aPPR and pPPR clusters. This module was therefore selected as a pan-placodal module.
510 GM2 is expressed in both neural crest and delaminating neural crest clusters (Figure 4; NC,
511 dNC), contains neural crest specifiers like *FoxD3*, *Snai2*, *Sox9*, *Pax7* and *Msx1*, as well as
512 the Wnt ligands *Wnt1* and *Wnt6* (Figure 4–Source Data 1) and thus represents a *bona fide*
513 neural crest specific module. Although GM1 contains multiple neural crest markers (*Ets1*,

514 *Sox10*, *Sox8*; Figure 4–Source Data 1), it is only expressed in delaminating neural crest cells
515 (Figure 4) and was therefore excluded for further analysis. Thus, GM5 and GM2 were
516 considered as pan-placodal and pan-neural crest module and used for co-expression analysis.
517



518

519 **Figure 4. Gene modules calculated at ss8. (A)** Heatmap showing gene modules calculated
520 from ss8 cells and then filtered to include only those that are differentially expressed between
521 cell states (see methods; full list of gene modules available in Figure 4-Source Data 1). GM5
522 and GM2 are highlighted as the selected pan-placodal and pan-neural crest gene modules
523 used for subsequent analysis in Figure 5.
524

525

526 *Placodal and neural crest bridge cells exhibit NPB-like co-expression*

527 To investigate the co-expression of these modules in NPB cells and their derivatives, we first
528 selected NPB (aNPB, pNPB), neural crest (NC, dNC) and PPR (aPPR, pPPR, PPR) cell states
529 from HH5-ss8 (Figure 5A, red cells; thereafter referred to as NPB subset). Next, we re-
530 clustered this subset of cells based on the expression of all genes (Figure 5B). UMAP analysis
531 suggested a bifurcation of neural crest and placodal fates taking place between HH7 and ss4,
532 with NPB cells located at the boundary between neural crest and aPPR cell clusters (Figure
533 5B, C).
534

535 To examine the relative expression of pan-placodal and pan-neural crest modules across the
536 entire cell subset, we ordered cells according to the expression ratio of both modules and

537 grouped them into 10 evenly sized bins (Figure 5D and Figure 5–figure supplement 1). Cells
538 at the extreme ends exclusively express the pan-neural crest (Bin 1) or the pan-placodal gene
539 module (Bin 10). As expected, these cells represent mostly delaminating neural crest or
540 placodal cells, respectively. However, cells in Bin 3 show equal expression levels of both
541 modules and contain a mixture of NPB, PPR and neural crest cell states (Figure 5D). We then
542 mapped Bin 3 cells back to the NPB subset UMAP (Figure 5B) and find that they indeed
543 correspond to NPB cells located specifically between PPR and neural crest cell clusters
544 (Figure 5E). This contrasts with Bin 1 and 10 cells, which map to the opposite ends of the
545 UMAP and their respective fates (Figure 5E). Thus, most cells expressing intermediate levels
546 of both neural crest and placodal gene modules are classified as NPB.
547
548 Bin 3 cells contain a mixture of cells collected from HH5-ss8. To evaluate which cell states
549 express similar levels of pan-neural crest and pan-placodal modules at different
550 developmental times, we mapped Bin 3 cells back to the UMAPs of individual stages (Figure
551 5G). At HH5 and HH6, neural crest cell states are not defined yet (Figure 5F colour coded
552 according to Figure 1B-D and Figure 2A-C), and Bin 3 cells consist of a mixture of aPPR and
553 pNPB cells (Figure 5F, G). At HH7, Bin 3 contains almost no PPR cells, instead these cells
554 are mostly neural crest and NPB (Figure 5F, G). At ss4, the number of cells expressing similar
555 levels of the pan-placodal and pan-neural crest modules decreases sharply. Although at this
556 stage a NPB state based on the expression of NPB specifiers cannot be classified, Bin 3 cells
557 are found at the boundary between neural crest and PPR clusters. In fact, they represent the
558 bridge cells predicted to be in an undetermined transcriptional state by RNA velocity (Figure
559 2G). Once placodal and neural crest lineages are largely segregated at ss8, only very few
560 scattered cells express similar levels of pan-neural crest and pan-placodal modules (Figure
561 5G).
562
563 The analysis above only focuses on cells of the NPB subset expressing pan-placodal and pan-
564 neural crest modules at a similar level, which are likely to be transcriptionally indeterminate
565 and not yet committed to either of the two fates. To identify cells with varying levels of co-
566 expression and those that may transit to either fate, we visualised cells that co-express these
567 modules at different levels. We calculated pan-neural crest and pan-placodal module co-
568 expression per developmental stage and filtered cells below a minimum threshold of 0.3
569 (Figure 5H). At HH5/6 almost all cells classified as NPB and placodal cells co-express both
570 modules above the threshold. However, this changes at HH7, when co-expressing cells are
571 largely restricted to the NPB and neural crest clusters (Figure 5H). By ss4, they are only found
572 in bridge cells between placodal and neural crest clusters (Figure 2G).
573

Given the different proportions of cells co-expressing these modules at each stage, we investigated whether the distributions in co-expression levels varied between them (Figure 5I). From the early to late stages, the distribution gradually changes from a normal distribution towards a negative binomial distribution. This suggests that at HH5-HH7 cells co-express pan-neural crest and pan-placodal modules at similar levels, while only a small proportion of co-expressing cells express these modules at high levels later.

Our analysis proposes that NPB cells are transcriptionally indeterminate, co-expressing gene modules that characterise both neural crest and placodal cells. This expression profile is likely to endow them with the potential to give rise to both, if not all NPB fates. The proportion of these cells decreases over time, but bridge cells at ss4 retain this characteristic and are therefore likely to retain bi- or multipotency even as the neural tube closes.

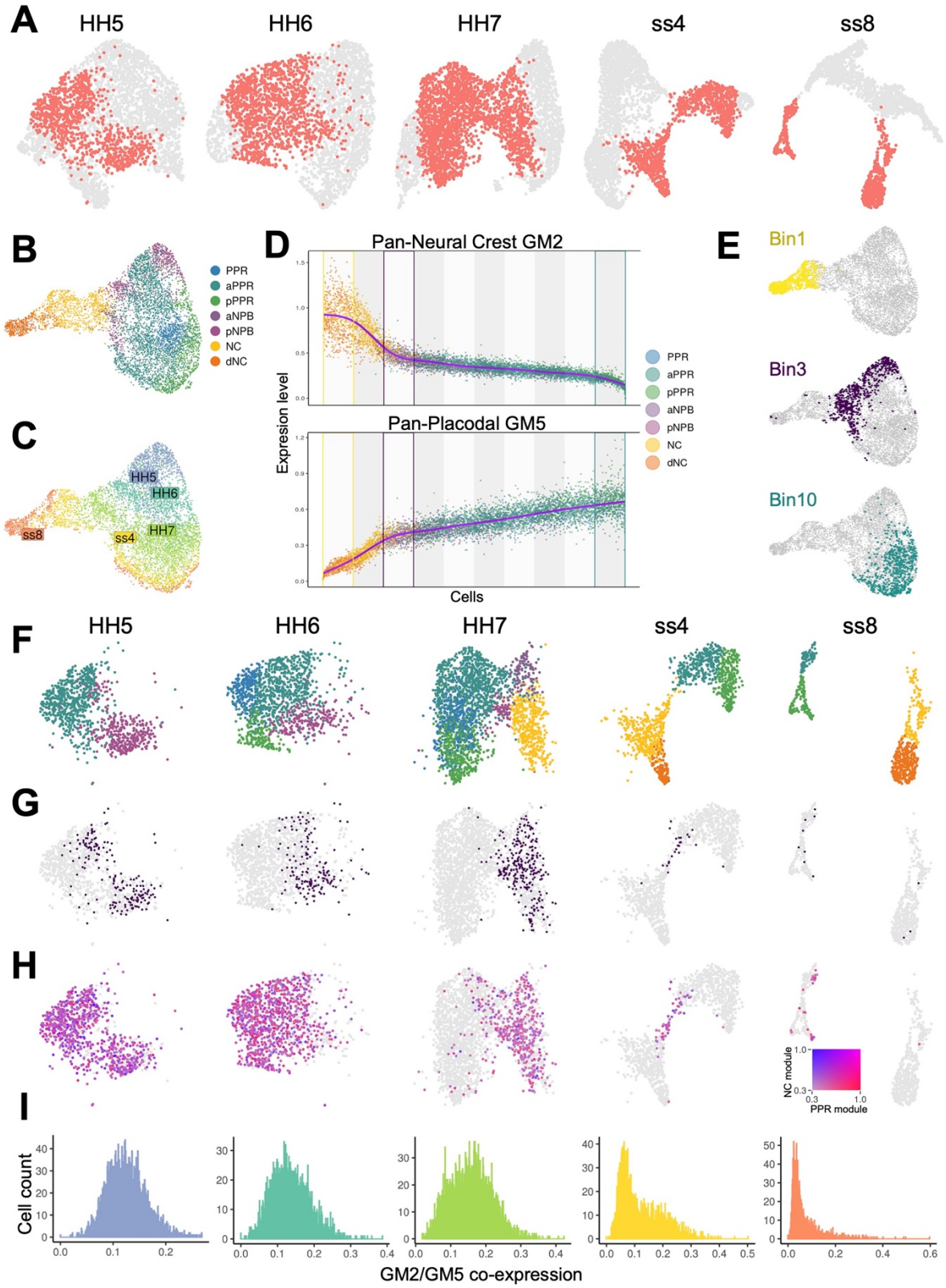


Figure 5. Cells at the NPB co-express placodal and neural crest gene modules. (A) UMAP plots of from each stage with all PPR, neural crest and NPB cells highlighted in red, which were selected as the 'NPB subset'. **(B)** UMAP plot of the NPB subset coloured by cell state. **(C)** UMAP plot of the NPB subset coloured by developmental stage. **(D)** Average

normalized expression of placodal and neural crest gene modules (GM5 and GM2, respectively) identified at ss8 (see Figure 4). Cells are ordered by their expression ratio of these two modules and coloured by cell state. Cells were grouped into 10 evenly sized bins (background shaded regions). Three bins (1, 3 and 10) are overlaid to highlight cells with high neural crest/low placodal (bin 1), equivalent levels of both (bin 3) and high placodal/low neural crest (bin 10) gene expression. **(E)** UMAP plots of the NPB subset displaying the cells captured within bins 1, 3 and 10. **(F-H)** UMAP plots from each developmental stage, displaying only cells in the NPB subset (equivalent to red cells in A). **(F)** Cells are coloured by cell state. **(G)** Cells which are captured in bin 3 are highlighted. **(H)** Co-expression analysis of the pan-neural crest and pan-placodal modules (see methods). Cells which co-express both modules above 0.3 are coloured. **(I)** Histograms revealing the distribution of co-expression of the pan-neural crest and pan-placodal modules at each developmental stage. At later stages (ss4-ss8), the distribution of co-expression shifts from a normal to a negative binomial distribution.

A temporal hierarchy of gene module expression in neural crest and placode precursors

To investigate the mechanism by which NPB cells segregate into the neural crest and placodal fates, we modelled gene module dynamics across the NPB subset alone (for workflow explanation see Figure 3–figure supplement 2). Previous modelling dynamics (Figure 3) analysed gene modules across the whole dataset to uncover sequential activation of placodal, neural and neural crest gene programmes at higher level. In contrast, focusing on the NPB subset provides a higher gene module resolution and consequently allows modelling of multiple placodal and neural crest gene modules and establishing the order in which different genes are activated.

As before, we first defined gene modules characteristic for neural crest and placodal cells. Gene modules were calculated using the neural crest, placodal and NPB cell state subset from HH5-ss8 as described above. Only 10 modules with differential expression between the neural crest and placodal fates were considered. These were filtered further to include only those that contained genes defined in the binary knowledge matrix as present in neural crest, delaminating neural crest or PPR cell states. This analysis identified three modules (GM12-14) upregulated in placodal and four in neural crest cells (GM40 and GM42-44) between HH7 and ss8 (Figure 6E, Figure 6–Source Data 1).

To predict cell trajectories towards neural crest or placodal terminal states prior to GM modelling, we performed scVelo analysis on the NPB subset (Figure 6B-D). As before, we used velocity and CellRank to calculate latent time and determine the probability that each cell

630 will become neural crest or placodal (Figure 6A-D). In turn, latent time and fate probabilities
631 were used to model gene module dynamics (Figure 6F-H).
632
633 Of the placodal modules, GM14 and GM12 are expressed as early as HH5. GM12 is
634 expressed throughout both aPPR and pNPB and contains early non-neural/PPR markers like
635 Gata2/3 (Figure 6E, Figure 6–Source Data 1). At HH6, GM12 is downregulated in the NPB
636 and remains confined to the PPR clusters until ss8. Unlike GM12, GM14 is initially restricted
637 to the PPR and only activated in the NPB at HH7. At ss4 and ss8 its expression is confined to
638 the PPR, with a higher expression in anterior cells. This module contains pan-placodal genes
639 like Dlx5/6 as well as anterior markers Six3 and Pax6 (Figure 6E, Figure 6–Source Data 1).
640 GM13 is the last placodal module to be activated and contains the definitive placodal
641 transcripts *Six1* and *Eya2*. It is first expressed at low levels at HH6 before being broadly
642 activated across the PPR at HH7 through to ss8. We also observe weaker expression of this
643 module within NPB cells at HH7 (Figure 6E). To visualise the sequential activation of these
644 modules across latent time, we used the same approach described above (Figure 3F-J),
645 whereby gene expression of each cell was weighted based on its placodal fate probability.
646 This analysis shows that GM12 is activated before GM14 and GM13 in latent time (Figure 6F).
647 Whilst the expression of GM12 and GM14 does not change dramatically, GM13 is highly
648 upregulated over latent time within the placodal cells (Figure 6F).
649
650 The four neural crest gene modules each contain *bona fide* neural crest transcriptional
651 regulators including *Msx1*, *Zic1*, *Zeb2*, *Lmx1b* (GM40), *Pax7*, *Snai2*, *Axud1*, *Foxd3*, *Tfap2b*,
652 *Sox9* (GM42), *Sox5*, *Id1/2* (GM44) and *Sox10*, *Ets1*, *Lmo4* and *Sox8* (GM43) (Figure 6E and
653 Figure 6–Source Data 1). Interestingly, our analysis groups these factors unbiasedly into four
654 modules, and visualisation of their expression across latent time weighted by neural crest fate
655 probability shows a clear sequential activation from GM40 to GM42, GM44 and GM43 (Figure
656 6G). GM40 is activated first, initially broadly across aPPR and pNPB clusters at HH5 and HH6,
657 before becoming restricted to neural crest and NPB clusters from HH7 onwards. By ss8, it is
658 downregulated in the delaminating neural crest population. GM42 and GM44 are highly
659 upregulated only in neural crest and delaminating neural crest cells, with GM42 first expressed
660 at HH7 and GM44 at ss4. Lastly, GM44 is activated specifically within the delaminating neural
661 crest cells at ss4 where it is further upregulated at ss8 (Figure 6E). This order suggests a
662 hierarchical relationship between these factors in the neural crest gene regulatory network.

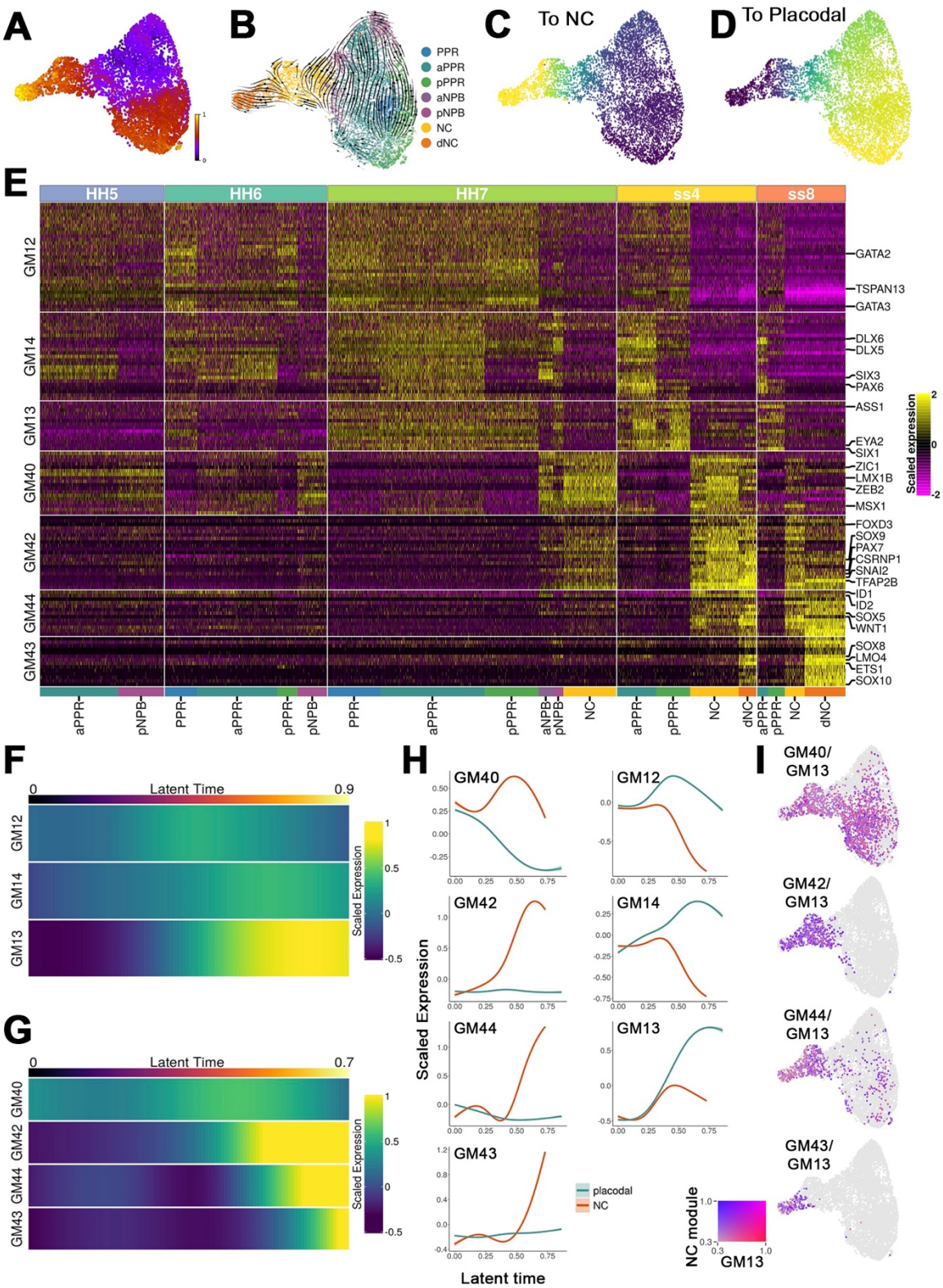


Figure 6. Co-expression analysis suggests neural crest specification requires downregulation of placodal genes. **(A)** UMAP plot of the NPB subset with cells coloured by latent time. **(B)** UMAP plot with cells coloured by cell state classifications and overlaid with RNA velocity vectors depicting the predicted directionality of transcriptional change. **(C-D)** UMAP plots showing the fate absorption probabilities of each cell towards the specified

670 terminal states (placodal or neural crest). **(E)** Heatmap showing gene modules calculated
671 across the NPB subset (full list of gene modules available in Figure 6-Source Data 1). To
672 identify gene modules potentially involved in placodal or neural crest specification, gene
673 modules were filtered to include those that show differential expression between placodal and
674 neural crest cell states (see methods). Genes specified as neural crest or placodal in our
675 binary knowledge matrix (Supplementary File 1) were selected as bait genes to further filter
676 the gene module list for visualisation. **(F)** Heatmap displaying predicted scaled expression of
677 placodal gene modules (GM12, GM13 and GM14) within the placodal lineage across latent
678 time. Gene expression was predicted by fitting a GAM which models scaled gene expression
679 as a function of latent time. Gene expression was weighted in the GAM using the placodal fate
680 lineage absorption probability. **(G)** Heatmap displaying predicted scaled expression of neural
681 crest gene modules (GM40, GM42, GM43 and GM44) within the neural crest lineage across
682 latent time. Gene expression was modelled as in (F) but using neural crest fate absorption
683 probabilities to weight gene expression. **(H)** Gene module dynamics plots for all 7 gene
684 modules: GM40, GM42, GM44, GM43 (neural crest) and GM12, GM14, GM13 (placodal).
685 Gene module dynamic plots display GAMs of scaled gene module expression across latent
686 time calculated as in (G) and (H). GAMs were weighted by the fate absorption probabilities of
687 each cell towards either placodal or neural crest terminal states.
688
689
690 While gene module selection was based on prior knowledge of marker genes, this analysis
691 also allows us to predict novel factors with similar expression profiles. These include *Nfkb1*,
692 *Ass1* and *Fam184b* in placodal cells and *Enc1*, *Rftn2* and *Olfml3* in neural crest clusters.
693 Modelling their expression across latent time reveals highly conserved dynamics with known
694 factors (Figure 6–figure supplement 2), thus raising the possibility that they may play a role in
695 placode and neural crest cell specification.
696
697 In summary, analysing gene module dynamics across the NPB subset identifies reveals the
698 sequential activation of different co-regulated genes in placodal and neural crest cell. Many
699 genes within these modules have been previously characterised and are known to form part
700 of the neural crest or placode gene regulatory network (for review: Grocott et al., 2012; Pla &
701 Monsoro-Burq, 2018; Schlosser, 2006, 2014; Simoes-Costa & Bronner, 2015; Thiery et al.,
702 2020). This unbiased approach groups these factors and predicts a hierarchical order, as well
703 as identifying new candidate factors which may play important roles in the segregation of the
704 neural crest and placodal fates (Figure 6–figure supplement 2).
705
706 **Emergence of neural crest cells may require the downregulation of definitive placodal**
707 **genes**
708
709 Various factors within the placodal and neural crest gene modules have previously been
710 implicated in conferring neural crest and placodal identity. To assess their temporal activation,

711 we modelled their dynamics across both developmental trajectories simultaneously and
712 assessed their co-expression in individual cells.
713
714 As with the previous analysis (Figure 3F-J) we found that gene modules varied in their
715 expression dynamics across both fates over latent time. Of the neural crest modules, GM40
716 (containing *Msx1* and *Zic1*) is the first to be activated in both the neural crest and placodal
717 cells before being downregulated in placode precursors at HH6/7 (Figure 6E, H). In contrast,
718 GM42, GM44 and GM43 are exclusively upregulated in the neural crest lineage (Figure 6H).
719 Of the placodal modules, GM12 and GM14 are initially highly expressed in both cell types and
720 then downregulated in the neural crest lineage. In contrast, GM13 is not initially expressed,
721 but starts to be upregulated in both neural crest and placodal cells as time progresses, before
722 being rapidly downregulated in neural crest cells as its expression increases in placode
723 precursors. Interestingly, GM13 contains definitive placodal markers *Six1* and *Eya2* which in
724 turn repress neural crest factors including *Pax7* (Brugmann et al., 2004; Christophorou et al.,
725 2009). Its initial upregulation within the neural crest cells suggests that their specification
726 requires the downregulation of this module. In contrast to GM13, placodal cells never
727 upregulate definitive neural crest modules (GM42-44).
728
729 Since neural crest precursors initially appear to express the late placodal module GM13, we
730 sought to investigate which cells co-express GM13 with each of the four neural crest modules.
731 As expected from the expression of the early neural crest module GM40 in placodal
732 precursors, we find broad co-expression of both modules (Figure 6I) within neural crest and
733 placodal cell clusters both, at the bifurcation point and in segregated clusters. In contrast, the
734 later neural crest modules are not co-expressed with GM13 extensively within the placodal
735 branch. However, we do find co-expression within the neural crest branch. This suggests that
736 some residual placodal gene expression remains in the neural crest clusters even at ss8. This
737 supports our previous finding that the specification of the neural crest state may require the
738 downregulation of placodal gene modules (Figure 3).
739

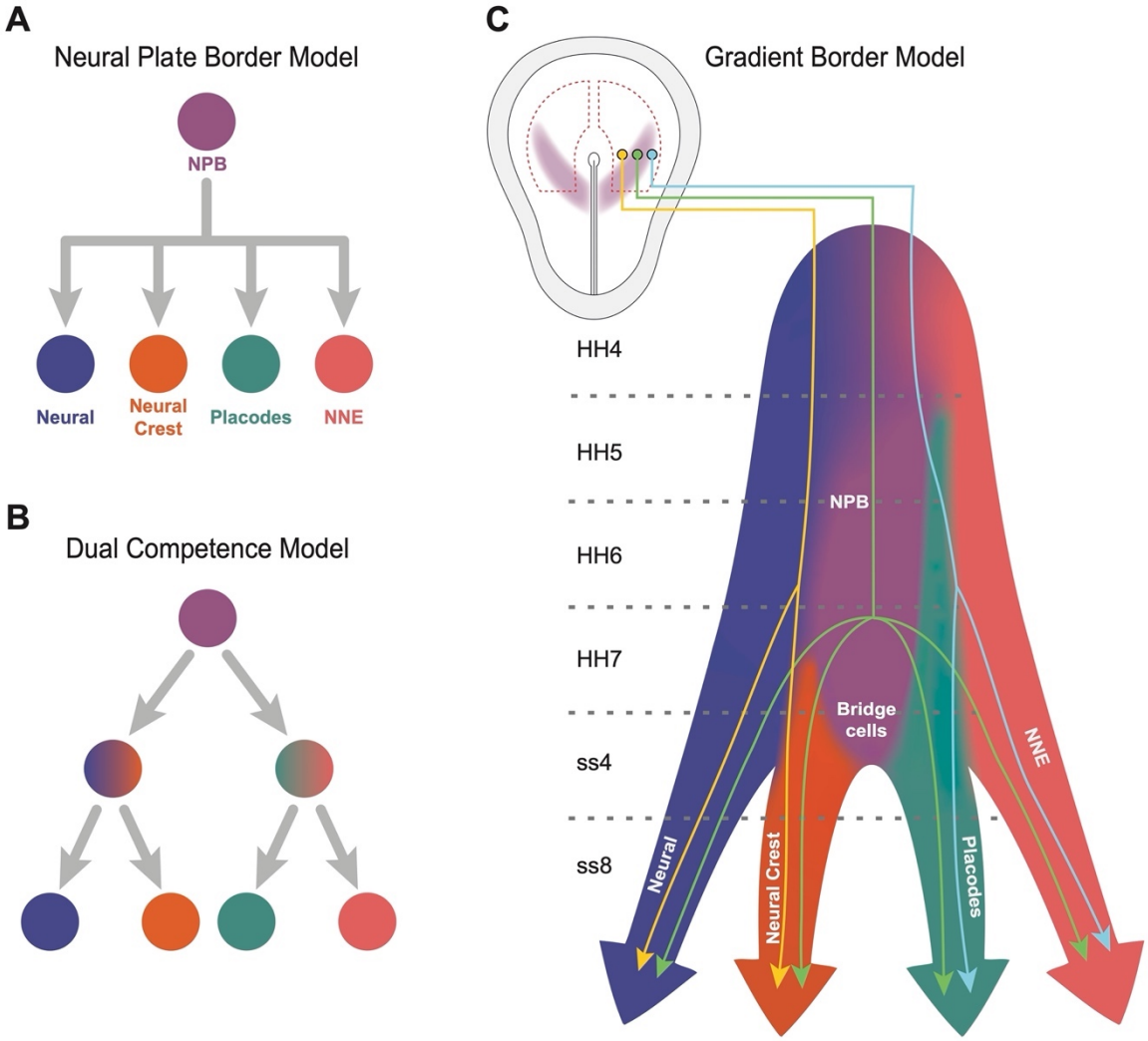


Figure 7. Gradient border model resolves conflicting models of lineage segregation at the neural plate border. (A) Schematic illustrating the neural plate border model. This model illustrates that cells at the neural plate border are multipotent and able to give rise to neural, neural crest, placodal and non-neural ectoderm (NNE) lineages. **(B)** Schematic illustrating the dual competence model. This model suggests that the neural plate border is first defined into non-neural/placodal and neural crest/neural competence domains prior to further subdivision into each of the final four cell lineages. (C) Schematic illustrating the gradient border model. This model proposes a probabilistic model of cell fate allocation which is intrinsically linked to the spatiotemporal positioning of a cell. We suggest that cells located within the medial neural plate border (yellow cell lineage) will give rise to neural and neural crest, cells within the lateral neural plate border (blue cell lineage) to non-neural ectoderm and placodes, whilst subsets of cells (including bridge cells) will remain in an indeterminate state and continue to give rise to all four lineages even at late stages of neurulation (green cell lineage). Shaded background colours represent the cell state space.

759 **Discussion**

760

761 The regulation of cell fate specification at the NPB has been extensively studied in a range of
762 model systems, giving rise to a wealth of knowledge regarding transcriptional regulators and
763 their interactions. Nevertheless, our understanding as to how individual cells undergo cell fate
764 decisions has been limited by our inability to study transcription and gene regulation at a single
765 cell level. Recent studies have demonstrated that in contrast to what was previously thought,
766 cells co-express markers that specify either placodal, neural or neural crest fates, even during
767 late neurulation (Roellig et al., 2017; Williams et al., 2022). This newly characterised cellular
768 heterogeneity requires us to revisit models for cell fate specification at the NPB and reconsider
769 our definition of the NPB cell state itself.

770

771 Our study unbiasedly characterised ectodermal cell states and cellular heterogeneity from
772 primitive streak stages though to late neurulation in chick. We demonstrate that although
773 significant transcriptional differences are apparent between cells at primitive streak stages,
774 due to high levels of heterogeneity, cells do not cluster into discrete cell states exhibiting
775 restricted expression patterns until the 1-somite stage. Instead, cell states are initially
776 characterised by the overlapping expression of broad early neural and non-neural genes, as
777 well as axial patterning markers. We identify increased cell state diversification during
778 neurulation, and segregation of neural, neural crest and placodal lineages taking place across
779 1-somite to 8-somite stages. High resolution transcriptional analysis of cells at the NPB reveals
780 that by the 4-somite stage, placodal and neural crest lineages have largely segregated.
781 However, despite this segregation, we continue to identify high levels of co-expression of
782 placodal and neural crest genes within bridge cells which are predicted to be in a dynamic
783 transcriptional state.

784

785 *Specification of neural and neural crest cell lineages requires downregulation of broad epiblast*
786 *modules*

787

788 Although specification networks for the neural, neural crest and placodal lineages are well
789 established (for review: Betancur et al., 2010; Grocott et al., 2012; Pla & Monsoro-Burq, 2018;
790 Stundl et al., 2021; Thiery et al., 2020), the order in which these lineages are specified remains
791 unclear. Our unbiased gene module analysis has identified groups of genes which correlate
792 and are specifically upregulated within each cell fate. Our gene modules support the
793 specification networks established in previous studies. *Pax7*, *Tfap2b*, *Snai2* and *Sox10* cluster
794 together and are upregulated in neural crest states, while *Six1*, *Eya2*, *Gata2/3* and *Dlx5/6* are
795 activated within placodal states and *Lmo1* and *Sox21* in neural states.

We also find putative novel regulators of these cell states, as well as the order in which these cell fates are specified. We note the sequential temporal activation of placodal, neural and neural crest specific gene modules, suggesting that these states are specified in that order. However, there is also asynchrony in the initiation of cell specification; although we identify placodal cells as early as HH5, we also observe indeterminate NPB cells which are only later specified. We suggest cell fate allocation at the NPB is a gradual process, with the NPB narrowing over time as cells within this population are specified.

Alongside characterising placodal, neural and neural crest gene modules, we observe the expression of two modules which exhibit initial broad pan-epiblast expression at primitive streak stages, but which are confined to placodal cell states at the 4 and 8-somite stages. Within these epiblast modules, unsurprisingly we note the expression of important placodal and non-neural genes *Tfap2c*, *Homer2* and the BMP inhibitor *Bambi*. Alongside these markers, we also find the pluripotency marker *Sall4* and *Tspan13*, with the latter shown to be positively regulated by Six1 (Mehdizadeh et al., 2021). A recent study in chick suggests that based on transcriptional profiling a 'pre-border' state can already be defined prior to gastrulation. In vitro experiments show that the pre-gastrula epiblast is specified as NPB; once cells have acquired a 'pre-border' state they can become neural and placodal precursors depending on the signalling input (Trevers et al., 2018). Given that the expression the epiblast modules becomes confined to the placodes, we speculate that the specification of the neural crest and neural states requires the downregulation of these modules.

Redefining the neural plate border based on the co-expression of competing gene modules

Members of Zic, Tfap2, Pax, Msx and Dlx transcription factor families have been termed NPB specifiers (for review: Stundl et al., 2021; Thiery et al., 2020). Tfap2a has been identified as a pioneer factor which triggers the expression of other NPB specifiers and is required but not sufficient for both neural crest and placodal cell fates (Bhat et al., 2013; Rothstein & Simoes-Costa, 2020). The cell fate switching role of Tfap2a depends upon its association with its co-factors, Tfap2b and Tfap2c. Tfap2c is necessary for the induction of NPB specifiers *Msx1*, *Pax7* and *Zic1*, whilst Tfap2b is later required for downstream neural crest specification (Rothstein & Simoes-Costa, 2020).

Gene module analysis across the neural plate border highlights the sequential activation and restriction of previously characterised NPB specifiers to neural crest cells. Early NPB specifiers *Msx1* and *Zic1* are initially broadly expressed across pNPB and aPPR cells at HH5,

833 but later become downregulated in the placodal lineage and upregulated in the neural crest.
834 The gene module containing *Pax7* is activated slightly later and contains other NPB specifiers
835 *Axud1* and *Tfap2b*, as well as definitive neural crest markers *FoxD3* and *Snai2*. Unlike the
836 *Msx1* module, which is initially upregulated in both aPPR and pNPB clusters at HH5, the *Pax7*
837 module is specifically restricted to the NPB and subsequent neural crest clusters (Figure 6F,
838 H, and Figure 6–figure supplement 1). This analysis reveals that despite these markers having
839 been previously characterised as NPB specifiers, their expression becomes restricted to the
840 neural crest. In contrast, *Dlx5/6* and *Tfap2a* are both found within a gene module that is later
841 restricted to the placodes (Figure 3A).
842
843 Although *Tfap2a* is grouped in the same module as placodal specifiers *Six1* and *Eya2*, it is
844 expressed in both placodal and neural crest cell states, even at later stages. This expression
845 reflects the key role that *Tfap2a* plays as a pioneer factor in both neural crest and placodal
846 specification, however its expression alone is not sufficient to classify the NPB state. At HH5,
847 although *Tfap2a* is expressed within the pNPB cluster alongside other NPB specifiers, it is
848 also expressed within aPPR cells which never appear to be exposed to NPB specifiers
849 including *Pax7*, *Pax3* and *Tfap2b* (Figure 1E).
850
851 To classify the NPB as a discrete state, NPB markers need to be uniformly upregulated in
852 NPB cells which give rise to neural, non-neural ectoderm, neural crest and PPR fates. Given
853 the non-uniform expression of previously characterised NPB specifiers (Basch et al., 2006;
854 Ezin et al., 2009; Hong & Saint-Jeannet, 2007; McLarren et al., 2003; Rothstein & Simoes-
855 Costa, 2020; Streit & Stern, 1999), and the expression of *Tfap2a* in both the NPB and early
856 aPPR cells which are unlikely to give rise to other fates, we propose that the NPB state cannot
857 be accurately classified using just these previously established genes. Instead, we suggest
858 that the NPB is better defined by the co-expression of markers regulating alternative cell fates.
859
860 *Gradient border model resolves the neural plate border and binary competence dilemma*
861
862 The 'binary competence' and 'neural plate border' models have provided useful frameworks
863 for understanding cell fate determination during neural, neural crest and placodal segregation
864 (Figure 7A, B). The binary competence model suggests that the ectoderm is subdivided into
865 non-neural/placodal and neural/neural crest domains (Pieper et al., 2012), whilst the neural
866 plate border model suggests cells at the border remain multi-potent and can differentiated into
867 all ectodermal fates (Roellig et al., 2017). Due to the limitations of studying cell fate
868 competence at a population level, transplantation experiments have failed to resolve these
869 two models.

870

871 Our co-expression analysis of pan-neural crest and pan-placodal gene modules found that at
872 early stages, most cells at the NPB co-express these modules at similar levels. At later stages,
873 we see a divergence in the level of co-expression between cell populations, with the highest
874 co-expressing cells located at the branching points between the neural crest and placodal
875 lineages (Figure 5H and I). In summary, we find that over time, the putative indeterminate cell
876 population reduces in size until it is restricted to only a small number of bridge cells which lie
877 at the boundary between placodal and neural crest cell clusters at the 4-somite stage. Cell
878 fate allocation within these bridge cells is likely to be driven via transcription factor cross-
879 repression, given the previously characterised cross-repressive interactions taking place
880 between placodal, neural crest and neural specifiers (Brugmann et al., 2004; Christophorou
881 et al., 2009; Schlosser et al., 2008; Wakamatsu et al., 2004).
882

883 Given these findings, we propose a 'gradient border' model for cell fate allocation (Figure 7C).
884 This model suggests that cell fate allocation is probabilistic and intimately linked to a cells
885 spatiotemporal positioning. In accordance with the binary competence model, cells located
886 within the lateral NPB will give rise to epidermal and placodal fates, whilst those found medially
887 give rise to neural and neural crest fates. However, instead of a binary segregation of the NPB
888 into neural/neural crest and non-neural/placodal domains, we suggest that a subset of cells at
889 the NPB are kept in an indeterminate state whilst they continue to co-express factors involved
890 in the specification of alternative lineages. Both our co-expression analysis and previous
891 documentation of heterogeneity at the NPB support this model.
892
893

894 **Materials and Methods**

895

896 **Data collection and alignment**

897

898 *Single cell dissociation and 10X single-cell mRNA sequencing*
899 Fertilised chicken eggs were incubated at 38°C for 24-32 hours depending upon the
900 developmental stage collected. Embryos were pinned and had the endoderm and mesoderm
901 removed using dispase (10mg/ml) before dissecting the ectodermal region of interest (Figure
902 1A). Multiple embryos from the same stage were pooled prior to dissociation. Samples were
903 then dissociated in 250μl pre-heated FACSmax cell dissociation solution (Amsbio, T200100)
904 with papain (30U/ml; Sigma, P3125) for 20 minutes at 37°C. Samples were gently pipetted
905 every 5 minutes to facilitate dissociation. After dissociation, 250μl resuspension solution was

906 added to prevent any further dissociation (HBSS; non-acetylated BSA (1mg/ml; Invitrogen,
907 10743447); HEPES (0.01M); non-essential amino acids (1x; Thermo Fisher Scientific,
908 11140050); rock inhibitor Y-27632 (10μM; Stemcell Technologies). Cells were passed through
909 a 20μm cell strainer (Miltenyi Biotech, 130-101-812) to remove any debris, before pelleting
910 and resuspending in 500μl resuspension buffer. To remove any dead or dying cells as well as
911 any remaining doublets, 1μl 0.1mg/ml DAPI was added to the cell suspension before FAC
912 sorting. Given the low number of cells obtained from each embryo, samples for each stage
913 were collected over multiple days and stored in 90% MeOH with 10% resuspension solution
914 after FAC sorting. Prior to sequencing, samples from the same stage were pooled and
915 resuspended in DPBS with 0.5% non-acetylated BSA and 0.5U/μl RNAse inhibitor (Roche
916 3335399001).
917
918 Cells underwent single-cell mRNA sequencing at the Francis Crick Institute, London. Library
919 preparation was carried out using the 10x 3' single-cell mRNA v3 kit (10x Genomics), before
920 sequencing on the HiSeq 4000 (Illumina) to a target depth of 50k reads/cell. Samples were
921 sequenced in two batches (batch 1: HH4, HH6, ss4, ss8; batch 2: HH5, HH6, HH7, ss4), with
922 two stages (HH6 and ss4) sequenced in both batches to allow for the correction of any
923 potential batch effects. Data from the first batch has previously been used for the
924 characterisation of neural induction (Trevers et al., 2021) (ArrayExpress accession number:
925 E-MTAB-10408).
926
927 *Read alignment*
928 A custom pipeline was developed in Nextflow (v20.07.1) (Di Tommaso et al., 2017) to handle
929 all GTF processing and read alignment whilst maintaining full reproducibility.
930
931 Prior to scRNAseq alignment, the GalGal6 ensembl 97 GTF file was modified to prefix genes
932 from the Z and W sex chromosomes with their respective chromosome IDs (Z- and W-).
933 Furthermore, we noticed several mitochondrial genes in the GTF were missing their
934 chromosome ID, in turn preventing accurately calculating cell percentage mitochondrial
935 content downstream; mitochondrial genes were therefore also prefixed in the GTF (MT-).
936 Finally, we annotated the key neural crest specifier Snai2 in the GTF
937 (ENSGALG00000030902). Reads were processed and aligned to GalGal6 using CellRanger
938 4.0.0 (10x Genomics).
939

940 The aligned reads from CellRanger (BAM) were re-counted using Velocyto (La Manno et al.,
941 2018) to generate spliced and unspliced counts for each gene, which are used in subsequent
942 RNA velocity analysis.
943
944
945 **Downstream pre-processing**
946
947 All downstream analysis was wrapped into a custom Nextflow pipeline to simplify re-running
948 the analysis by handling parallel processing of data subsets.
949
950 *Data filtering*
951 Quality control, data filtering, cell clustering, dimensionality reduction and initial data
952 visualisation was carried out using Seurat v4.0.5 (Hao et al., 2021). Data filtering was carried
953 out on each sequencing batch separately before data integration. To ensure the removal of
954 poor-quality cells whilst also minimising over-filtering, we first applied a medium filtering
955 threshold (cells expressing fewer than 1000 genes, more than 6500 genes, or had more than
956 15% mitochondrial content were excluded from subsequent analysis) followed by the removal
957 of remaining poor-quality cell clusters.
958
959 *Dimensionality reduction and clustering*
960 Linear dimensionality reduction was carried out using principal component analysis (PCA).
961 The number of principal components (PCs) used to construct the subsequent shared-nearest-
962 neighbour (SNN) graph was determined unbiasedly, using an approach developed by the
963 Harvard Chan Bioinformatics Core (HBC). First, we calculate the PC at which subsequent PCs
964 contribute less than 5% of standard deviation, and the point where the cumulative variation of
965 PCs exceeds 90%. The larger of these two values is selected. Second, we calculate the PC
966 where the inclusion of consecutive PCs contributes less than 0.1% additional variation. The
967 smaller value from these two steps is selected as a cut-off.
968
969 Louvain clustering was carried out with varying resolutions depending upon the data subset
970 in question. A high clustering resolution (resolution=2) was used to identify poor quality
971 clusters and contaminating cell states. This was to ensure that only poor-quality cells were
972 filtered from the dataset. All other clustering was carried out using a resolution of 0.5.
973
974 *Filtering of poor-quality clusters*

975 Cell quality was measured using unique gene and UMI counts. Clusters in which both metrics
976 fell below the 25th percentile relative to the rest of the dataset were considered poor quality
977 and excluded from further analysis.
978
979 *Data integration*
980 After the removal of poor-quality cells, cells from the two separate batches were integrated
981 into a single full dataset (Figure 1–figure supplement 1H-I). To avoid over-correction and
982 removal of stage differences, we integrated the data from the two sequencing batches using
983 STACAS v1.1.0 (Andreatta & Carmona, 2021). This R package allows for alignment of
984 datasets containing partially overlapping cell states.
985
986 *Regressing out confounding variables*
987 Before downstream analysis, we regressed out percentage mitochondrial content, cell sex,
988 and cell cycle. After regressing out percentage mitochondrial content, we identified a residual
989 sex effect whereby cells clustered according to the expression of either W or Z chromosome
990 genes. This resulted in a duplication of cell clusters, most apparent at HH4 (Figure 1–figure
991 supplement 1A, B). Cells were k-means clustered according to their expression of W
992 chromosome genes and classified as either male or female. This classification was used to
993 regress out cell sex. Cell cycle was observed to have a strong effect on cell clustering and
994 was therefore also regressed out during data scaling (Figure 1–figure supplement 1C, D).
995
996 *Removal of contaminating clusters*
997 We found small contaminating clusters of primordial germ cells, blood islands, mesoderm and
998 endoderm within the dataset (Figure 1–figure supplement 1F, G). To semi-unbiasedly remove
999 these clusters, we calculated the average expression of candidate markers for each of these
1000 cell states and removed clusters in which the median expression of any of these modules was
1001 greater than the 90th percentile relative to the rest of the dataset (Figure 1–figure supplement
1002 1F'', G).
1003
1004 **Downstream analysis**
1005
1006 *Unbiased cell classification*
1007 Cell states were classified on each stage independently using a binary knowledge matrix
1008 (Supplementary File 1). First, the full dataset was split by stage and then re-scaled and re-
1009 clustered as described above. Our binary knowledge matrix consists of binarized expression
1010 of 76 genes across 24 cell states. For each gene, published *in situ* hybridisation expression
1011 patterns were used to determine expression within each defined cell state. Given that gene

1012 expression can vary dramatically between stages, and some cell states are known to not be
1013 present at specific timepoints (i.e., delaminating neural crest at HH4), we restricted the
1014 possible cell states at specific stages. At HH4 there were 7 possible cell states, including non-
1015 neural ectoderm (NNE), node, streak, extra-embryonic (EE), early NPB (eNPB), early neural
1016 (eN), early caudal neural (eCN). At HH5 there were 17 potential states, this included all cell
1017 states in the binary knowledge matrix, except for the later counterparts to the earlier cell states
1018 which are ventral forebrain (vFB), forebrain (FB), midbrain (MB), hindbrain (HB), neural crest
1019 (NC), delaminating neural crest (dNC) and presumptive epidermis (pEpi). At HH6 we included
1020 the same cell states as HH5 except for EE and eNPB given that no cell clusters were classified
1021 as such at HH5. For stages HH7, ss4 and ss8 we included 17 cell states, excluding the
1022 following states which are only found at earlier stages (NNE, node, streak, eN, eCN, eNPB,
1023 EE).
1024
1025 For each cell, we calculated the average scaled expression of the genes expressed in each
1026 cell state according to our binary knowledge matrix. Cell clusters were then assigned a cell
1027 state, depending on which cell state exhibited the highest median expression in that given cell
1028 cluster. Assigned cell states at each of the 6 sampled developmental stages were then
1029 transferred to the full dataset.
1030
1031 *Gene module identification and filtering*
1032 To calculate groups of highly correlated genes, we first filtered genes which do not correlate
1033 (Spearman correlation < 0.3) with at least 3 other genes. Then we calculated gene modules
1034 using the Antler R package (Delile et al., 2019), which iteratively clusters genes hierarchically
1035 based on their gene-gene Spearman correlation. After each iteration, gene modules were
1036 filtered to remove gene modules which are not highly expressed (expressed in fewer than 10
1037 cells) or which do not show consistent high expression across genes (fewer than 40% of genes
1038 in the module are expressed after binarization). Gene modules were calculated on different
1039 subsets of cells: the full dataset (Figure 3 and Figure 3–figure supplement 1), ss8 (Figure 4
1040 and Figure 5), and the NPB subset (Figure 6). Calculating gene modules on different subsets
1041 allowed us to zoom in on specific aspects of ectodermal lineage segregation (for workflow
1042 explanation see Figure 3–figure supplement 2).
1043
1044 Many genes correlate across the dataset, but do not show differential expression between cell
1045 states and therefore do not provide useful information regarding lineage segregation.
1046 Therefore, to identify which gene modules correlate with lineage segregation, modules were
1047 filtered to only include those in which more than 50% of their genes were differentially
1048 expressed (logFC > 0.5, adjusted p-value < 0.001) in at least one cell state relative to the rest

of the dataset. For the NPB subset, this filtering of differential expression was just between the neural crest (NC, dNC) and placodal (aPPR, pPPR, PPR) states rather than between all cell states (logFC > 0.25, adjusted p-value < 0.001).

To remove gene modules which correlate with technical variation, we filtered those in which more than 50% their genes showed differential expression between sequencing batches (logFC > 0.25, adjusted p-value <0.001).

For the gene modules calculated on the full dataset, we were particularly interested in gene modules that segregated into one of the three lineages (neural, neural crest or placodal) at later stages. To focus our analysis on these gene modules, we applied an additional filtering step to only keep gene modules in which more than 50% of genes were differentially expressed between one of the three lineages at ss8 (logFC > 0.25, adjusted p-value <0.001). After filtering 9 gene modules remained (Figure 3–figure supplement 2A, Figure 3–Source Data 1). Modules were further manually filtered for visualisation (Figure 3A).

For the gene modules calculated on the NPB subset we wanted to focus on the co-expression and sequential activation of gene modules that contain well-characterised neural crest and PPR markers. We therefore further filtered NPB gene modules to include only those that had genes that were defined within neural crest, delaminating neural crest or PPR cell states in the binary knowledge matrix. After this final filtering step, 7 gene modules remained (Figure 6E, Figure 6–Source Data 1).

Identification of undetermined cells at the NPB

To identify which cells expressed competing pan-placodal and pan-neural crest modules at similar levels, we calculated the average expression (normalised counts) of these two gene modules in each cell and then ordered cells by their ratio of pan-placodal to pan-neural crest module expression. Cells were then grouped into 10 evenly sized bins and visualised on UMAPs.

Co-expression distribution plots

To investigate the distribution of co-expression between cells, co-expression histograms were plotted (Figure 5I). Co-expression values were calculated by multiplying the average normalised expression of the pan-placodal and pan-neural crest gene modules.

Co-expression visualisation

1085 The average normalised expression of each gene module was scaled between 0 and 1. A
1086 100x100 two-colour matrix was created using Seurat's BlendMatrix function. Each cell was
1087 assigned a colour from this matrix based on their scaled average expression of the gene
1088 modules (see Figure 5H and Figure 6I legends). Only cells which express both modules above
1089 a threshold of 0.3 after scaling were visualised.
1090
1091 *RNA velocity*
1092 Spliced and unspliced matrices obtained from Velocyto (La Manno et al., 2018) (see *Read*
1093 *Alignment*) were analysed using scVelo (Bergen et al., 2020) to generate cell velocity vectors.
1094 Velocyto matrices were intersected with each of the data subsets generated in Seurat. First
1095 and second order moments were then calculated for each cell across its 20 nearest
1096 neighbours using the top 20 PCs. RNA velocity was ran using the dynamical model whilst
1097 accounting for differential kinetics; this allows for differential splicing rates between genes as
1098 well as differential splicing kinetics between cell states.
1099
1100 Latent time was calculated and used downstream as a measure of a cell's internal clock. When
1101 calculating latent time, cells were rooted to the earliest timepoint present in the data subset;
1102 in cases where only one stage was present, the cell root was calculated automatically. For all
1103 latent time estimates, we used a diffusion weighting of 0.2.
1104
1105 *Modelling gene expression dynamics*
1106 Lineage inference was calculated using CellRank, which uses a probabilistic approach to
1107 assign cell lineage (Lange et al., 2022). Cell transition matrices were calculated using RNA
1108 velocity (0.8 weighting) and nearest neighbour graph (0.2 weighting) kernels. A softmax scale
1109 of 4 was set to prevent transitions tending towards orthogonality. When predicting cell
1110 transitions, terminal states were specified according to the lineages which we were wanting to
1111 model gene dynamics across. For the full dataset, we predicted cell transitions towards either
1112 neural, placodal or neural crest lineages. Terminal states were specified as: delaminating
1113 neural crest and neural crest cell states at ss4 and ss8 for the neural crest lineage; hindbrain,
1114 midbrain, and forebrain cell states at ss4 and ss8 for the neural lineage; anterior PPR,
1115 posterior PPR and PPR cell states at ss4 and ss8 for the placodal lineage (Figure 3–figure
1116 supplement 1C). For our NPB subset we predicted cell transitions towards either placodal or
1117 neural crest lineages, with terminal states specified as: delaminating neural crest and neural
1118 crest cell states at ss8 for the neural crest lineage; aPPR and PPR cell states at ss8 for the
1119 placodal lineage.
1120

Gene expression dynamics were calculated by fitting general additive models (GAM), which modelled scaled gene expression as a function of latent time using the mgcv package in R. Lineage absorption probability obtained from CellRank were used to weight a GAM for each of the predicted lineages. To generate a smoothed prediction of gene expression change, a cubic spline was fitted with knots=4. When modelling the dynamics of entire gene modules, all genes within a given module were included in the GAM.

Data availability and reproducibility

10x single cell RNAseq was carried out in two batches and is available under two separate accession numbers (ArrayExpress: E-MTAB-10408 and E-MTAB-1144). Our NGS alignments and downstream analysis have been wrapped into custom Nextflow pipelines allowing for full reproducibility. For the code used in this analysis, including links to our Docker containers, see our GitHub repository at https://github.com/alexthiery/10x_neural_plate_border.

Acknowledgements

We thank Teresa Rayon for assistance with establishing our single cell dissociation protocol and the Advanced Sequencing Facility at the Francis Crick Institute for the single cell RNA sequencing. We would also like to thank Igor Adameyko and Artem Artemov for guidance on co-expression analysis, Chantal Hubens for technical support, Alessandra Vigilante for comments on the manuscript, and Streit and Luscombe lab groups for helpful discussions. This work was supported by the Wellcome Trust (108874/B/15/Z) and the BBSRC (BB/S005536/1; BB/R006342/1).

Author contributions

A.T., A.S., J.B. designed study. A.T., A.B., E.H., A.S. collected samples and processed for scRNAseq. A.T., C.C. developed analysis pipeline and carried out data alignment. A.T., E.H. carried out downstream computational analysis of scRNAseq data. A.T., A.B., E.H., A.S. made figures and wrote manuscript. All authors provided feedback on the manuscript.

References

Adams, M. S., Gammill, L. S., & Bronner-Fraser, M. (2008). Discovery of transcription factors and other candidate regulators of neural crest development. *Dev Dyn, 237*(4), 1021-1033. doi:10.1002/dvdy.21513

Ahrens, K., & Schlosser, G. (2005). Tissues and signals involved in the induction of placodal Six1 expression in Xenopus laevis. *Dev Biol, 288*(1), 40-59.

Andreatta, M., & Carmona, S. J. (2021). STACAS: Sub-Type Anchor Correction for Alignment in Seurat to integrate single-cell RNA-seq data. *Bioinformatics, 37*(6), 882-884. doi:10.1093/bioinformatics/btaa755

Baker, C. V., & Bronner-Fraser, M. (1997). The origins of the neural crest. Part I: embryonic induction. *Mech Dev, 69*(1-2), 3-11. doi:10.1016/s0925-4773(97)00132-9

Baker, C. V., & Bronner-Fraser, M. (2001). Vertebrate cranial placodes I. Embryonic induction. *Dev Biol, 232*(1), 1-61. doi:10.1006/dbio.2001.0156

Barembaum, M., & Bronner-Fraser, M. (2007). Spalt4 mediates invagination and otic placode gene expression in cranial ectoderm. *Development, 134*(21), 3805-3814. doi:10.1242/dev.02885

Barth, K. A., Kishimoto, Y., Rohr, K. B., Seydler, C., Schulte-Merker, S., & Wilson, S. W. (1999). Bmp activity establishes a gradient of positional information throughout the entire neural plate. *Development, 126*(22), 4977-4987.

Basch, M. L., Bronner-Fraser, M., & Garcia-Castro, M. I. (2006). Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature, 441*(7090), 218-222.

Bell, G. W., Yatskievych, T. A., & Antin, P. B. (2004). GEISHA, a whole-mount in situ hybridization gene expression screen in chicken embryos. *Dev Dyn, 229*(3), 677-687. doi:10.1002/dvdy.10503

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol, 38*(12), 1408-1414. doi:10.1038/s41587-020-0591-3

Betancur, P., Bronner-Fraser, M., & Sauka-Spengler, T. (2010). Assembling neural crest regulatory circuits into a gene regulatory network. *Annu Rev Cell Dev Biol, 26*, 581-603. doi:10.1146/annurev.cellbio.042308.113245

Bhat, N., Kwon, H. J., & Riley, B. B. (2013). A gene network that coordinates preplacodal competence and neural crest specification in zebrafish. *Dev Biol, 373*(1), 107-117. doi:10.1016/j.ydbio.2012.10.012

Bhattacharyya, S., Bailey, A. P., Bronner-Fraser, M., & Streit, A. (2004). Segregation of lens and olfactory precursors from a common territory: cell sorting and reciprocity of Dlx5 and Pax6 expression. *Dev Biol, 271*(2), 403-414. doi:10.1016/j.ydbio.2004.04.010

Brugmann, S. A., Pandur, P. D., Kenyon, K. L., Pignoni, F., & Moody, S. A. (2004). Six1 promotes a placodal fate within the lateral neurogenic ectoderm by functioning as both a transcriptional activator and repressor. *Development, 131*(23), 5871-5881.

Buitrago-Delgado, E., Nordin, K., Rao, A., Geary, L., & LaBonne, C. (2015). Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science, 348*(6241), 1332-1335. doi:10.1126/science.aaa3655

Chapman, S. C., Schubert, F. R., Schoenwolf, G. C., & Lumsden, A. (2002). Analysis of spatial and temporal gene expression patterns in blastula and gastrula stage chick embryos. *Dev Biol, 245*(1), 187-199. doi:10.1006/dbio.2002.0641

Chen, B., Kim, E.-H., & Xu, P.-X. (2009). Initiation of olfactory placode development and neurogenesis is blocked in mice lacking both Six1 and Six4. *Developmental Biology, 326*(1), 75-85.

Christophorou, N. A., Bailey, A. P., Hanson, S., & Streit, A. (2009). Activation of Six1 target genes is required for sensory placode formation. *Dev Biol, 336*(2), 327-336.

Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., & Sagner, A. (2019). Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development, 146*(12). doi:10.1242/dev.173807

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nat Biotechnol, 35*(4), 316-319. doi:10.1038/nbt.3820

Ezin, A. M., Fraser, S. E., & Bronner-Fraser, M. (2009). Fate map and morphogenesis of presumptive neural crest and dorsal neural tube. *Dev Biol, 330*(2), 221-236. doi:10.1016/j.ydbio.2009.03.018

Garcia-Castro, M. I., Marcelle, C., & Bronner-Fraser, M. (2002). Ectodermal Wnt function as a neural crest inducer. *Science, 297*(5582), 848-851.

Grocott, T., Tambalo, M., & Streit, A. (2012). The peripheral sensory nervous system in the vertebrate head: a gene regulatory perspective. *Dev Biol, 370*(1), 3-23. doi:10.1016/j.ydbio.2012.06.028

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., III., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., Rogers, A. J., McElrath, J. M., Blish, C. A., Gottardo, R., Smibert, P., & Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell, 184*(13), 3573-3587 e3529. doi:10.1016/j.cell.2021.04.048

Heeg-Truesdell, E., & LaBonne, C. (2006). Neural induction in Xenopus requires inhibition of Wnt-beta-catenin signaling. *Developmental Biology, 298*(1), 71-86.

Hemmati-Brivanlou, A., Kelly, O. G., & Melton, D. A. (1994). Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell, 77*(2), 283-295.

Hidalgo-Sanchez, M., Millet, S., Bloch-Gallego, E., & Alvarado-Mallart, R. M. (2005). Specification of the meso-isthmo-cerebellar region: the Otx2/Gbx2 boundary. *Brain Res Brain Res Rev, 49*(2), 134-149. doi:10.1016/j.brainresrev.2005.01.010

Hidalgo-Sanchez, M., Millet, S., Simeone, A., & Alvarado-Mallart, R. M. (1999). Comparative analysis of Otx2, Gbx2, Pax2, Fgf8 and Wnt1 gene expressions during the formation of the chick midbrain/hindbrain domain. *Mech Dev, 81*(1-2), 175-178. doi:10.1016/s0925-4773(98)00224-x

Hintze, M., Prajapati, R. S., Tambalo, M., Christophorou, N. A. D., Anwar, M., Grocott, T., & Streit, A. (2017). Cell interactions, signals and transcriptional hierarchy governing placode progenitor induction. *Development, 144*(15), 2810-2823. doi:10.1242/dev.147942

Hollyday, M., McMahon, J. A., & McMahon, A. P. (1995). Wnt expression patterns in chick embryo nervous system. *Mech Dev, 52*(1), 9-25. doi:10.1016/0925-4773(95)00385-e

Hong, C. S., & Saint-Jeannet, J. P. (2007). The activity of Pax3 and Zic1 regulates three distinct cell fates at the neural plate border. *Mol Biol Cell, 18*(6), 2192-2202. doi:10.1091/mbc.e06-11-1047

Kwon, H. J., Bhat, N., Sweet, E. M., Cornell, R. A., & Riley, B. B. (2010). Identification of early requirements for preplacodal ectoderm and sensory organ development. *PLoS Genet, 6*(9). doi:10.1371/journal.pgen.1001133

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lonnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundstrom, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., & Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature, 560*(7719), 494-498. doi:10.1038/s41586-018-0414-6

Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., Lickert, H., Ansari, M., Schniering, J., Schiller, H. B., Pe'er, D., & Theis, F. J. (2022). CellRank for directed single-cell fate mapping. *Nat Methods*. doi:10.1038/s41592-021-01346-6

Liberatore, C. M., Searcy-Schrick, R. D., & Yutzey, K. E. (2000). Ventricular expression of tbx5 inhibits normal heart chamber development. *Dev Biol, 223*(1), 169-180. doi:10.1006/dbio.2000.9748

Liem, K. F., Jr., Tremml, G., Roelink, H., & Jessell, T. M. (1995). Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell, 82*(6), 969-979.

Linker, C., & Stern, C. D. (2004). Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. *Development, 131*(22), 5671-5681.

Litsiou, A., Hanson, S., & Streit, A. (2005). A balance of FGF, Wnt and BMP signalling positions the future placode territory in the head. *Development, 132*, 4051-4062.

1256 Londin, E. R., Niemiec, J., & Sirotkin, H. I. (2005). Chordin, FGF signalling, and mesodermal factors
1257 cooperate in zebrafish neural induction. *Dev Biol, 279*(1), 1-19.

1258 Marchant, L., Linker, C., Ruiz, P., Guerrero, N., & Mayor, R. (1998). The inductive properties of
1259 mesoderm suggest that the neural crest cells are specified by a BMP gradient. *Dev Biol, 198*(2),
1260 319-329.

1261 McLarren, K. W., Litsiou, A., & Streit, A. (2003). DLX5 positions the neural crest and preplacode region
1262 at the border of the neural plate. *Dev Biol, 259*(1), 34-47. doi:10.1016/s0012-1606(03)00177-5

1263 Mehdizadeh, T., Majumdar, H. D., Ahsan, S., Tavares, A. L. P., & Moody, S. A. (2021). Mutations in
1264 SIX1 Associated with Branchio-oto-Renal Syndrome (BOR) Differentially Affect Otic Expression
1265 of Putative Target Genes. *J Dev Biol, 9*(3). doi:10.3390/jdb9030025

1266 Milet, C., Maczkowiak, F., Roche, D. D., & Monsoro-Burq, A. H. (2013). Pax3 and Zic1 drive induction
1267 and differentiation of multipotent, migratory, and functional neural crest in Xenopus embryos.
1268 *Proc Natl Acad Sci U S A, 110*(14), 5528-5533. doi:10.1073/pnas.1219124110

1269 Monsoro-Burq, A. H. (2015). PAX transcription factors in neural crest development. *Semin Cell Dev*
1270 *Biol, 44*, 87-96. doi:10.1016/j.semcdb.2015.09.015

1271 Monsoro-Burq, A. H., Fletcher, R. B., & Harland, R. M. (2003). Neural crest induction by paraxial
1272 mesoderm in Xenopus embryos requires FGF signals. *Development, 130*(14), 3111-3124.

1273 Ohuchi, H., Tomonari, S., Itoh, H., Mikawa, T., & Noji, S. (1999). Identification of chick rax/rx genes with
1274 overlapping patterns of expression during early eye and brain development. *Mech Dev, 85*(1-
1275 2), 193-195. doi:10.1016/s0925-4773(99)00094-5

1276 Ota, K., Nagai, H., & Sheng, G. (2007). Expression and hypoxic regulation of hif1alpha and hif2alpha
1277 during early blood and endothelial cell differentiation in chick. *Gene Expr Patterns, 7*(7), 761-
1278 766. doi:10.1016/j.modgep.2007.05.007

1279 Ozaki, H., Nakamura, K., Funahashi, J., Ikeda, K., Yamada, G., Tokano, H., Okamura, H. O., Kitamura,
1280 K., Muto, S., Kotaki, H., Sudo, K., Horai, R., Iwakura, Y., & Kawakami, K. (2004). Six1 controls
1281 patterning of the mouse otic vesicle. *Development, 131*(3), 551-562.

1282 Patthey, C., Gunhaga, L., & Edlund, T. (2008). Early development of the central and peripheral nervous
1283 systems is coordinated by Wnt and BMP signals. *PLoS ONE, 3*(2), e1625.
1284 doi:10.1371/journal.pone.0001625

1285 Patthey, C., Schlosser, G., & Shimeld, S. M. (2014). The evolutionary history of vertebrate cranial
1286 placodes--I: cell type evolution. *Dev Biol, 389*(1), 82-97. doi:10.1016/j.ydbio.2014.01.017

1287 Paxton, C. N., Bleyl, S. B., Chapman, S. C., & Schoenwolf, G. C. (2010). Identification of differentially
1288 expressed genes in early inner ear development. *Gene Expr Patterns, 10*(1), 31-43.
1289 doi:10.1016/j.gep.2009.11.002

1290 Pera, E., Stein, S., & Kessel, M. (1999). Ectodermal patterning in the avian embryo: epidermis versus
1291 neural plate. *Development, 126*(1), 63-73.

1292 Pieper, M., Ahrens, K., Rink, E., Peter, A., & Schlosser, G. (2012). Differential distribution of
1293 competence for panplacodal and neural crest induction to non-neural and neural ectoderm.
1294 *Development, 139*(6), 1175-1187. doi:10.1242/dev.074468

1295 Pla, P., & Monsoro-Burq, A. H. (2018). The neural border: Induction, specification and maturation of the
1296 territory that generates neural crest cells. *Dev Biol, 444 Suppl 1*, S36-S46.
1297 doi:10.1016/j.ydbio.2018.05.018

1298 Plouhinec, J. L., Roche, D. D., Pegoraro, C., Figueiredo, A. L., Maczkowiak, F., Brunet, L. J., Milet, C.,
1299 Vert, J. P., Pollet, N., Harland, R. M., & Monsoro-Burq, A. H. (2014). Pax3 and Zic1 trigger the
1300 early neural crest gene regulatory network by the direct activation of multiple key neural crest
1301 specifiers. *Dev Biol, 386*(2), 461-472. doi:10.1016/j.ydbio.2013.12.010

1302 Prasad, M. S., Charney, R. M., & Garcia-Castro, M. I. (2019). Specification and formation of the neural
1303 crest: Perspectives on lineage segregation. *Genesis, 57*(1), e23276. doi:10.1002/dvg.23276

1304 Prasad, M. S., Sauka-Spengler, T., & LaBonne, C. (2012). Induction of the neural crest state: control of
1305 stem cell attributes by gene regulatory, post-transcriptional and epigenetic interactions. *Dev*
1306 *Biol, 366*(1), 10-21. doi:10.1016/j.ydbio.2012.03.014

Reichert, S., Randall, R. A., & Hill, C. S. (2013). A BMP regulatory network controls ectodermal cell fate decisions at the neural plate border. *Development, 140*(21), 4435-4444. doi:10.1242/dev.098707

Rex, M., Orme, A., Uwanogho, D., Tointon, K., Wigmore, P. M., Sharpe, P. T., & Scotting, P. J. (1997). Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. *Dev Dyn, 209*(3), 323-332. doi:10.1002/(SICI)1097-0177(199707)209:3<323::AID-AJA7>3.0.CO;2-K

Roellig, D., Tan-Cabugao, J., Esaian, S., & Bronner, M. E. (2017). Dynamic transcriptional signature and cell fate analysis reveals plasticity of individual neural plate border cells. *Elife, 6*. doi:10.7554/eLife.21620

Rothstein, M., & Simoes-Costa, M. (2020). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res, 30*(1), 35-48. doi:10.1101/gr.249680.119

Sasai, Y., Lu, B., Steinbeisser, H., & De Robertis, E. M. (1995). Regulation of neural induction by the Chd and Bmp-4 antagonistic patterning signals in Xenopus. *Nature, 376*(6538), 333-336.

Schlosser, G. (2006). Induction and specification of cranial placodes. *Dev Biol, 294*(2), 303-351. doi:10.1016/j.ydbio.2006.03.009

Schlosser, G. (2014). Early embryonic specification of vertebrate cranial placodes. *Wiley Interdiscip Rev Dev Biol, 3*(5), 349-363. doi:10.1002/wdev.142

Schlosser, G., Awtry, T., Brugmann, S. A., Jensen, E. D., Neilson, K., Ruan, G., Stammler, A., Voelker, D., Yan, B., Zhang, C., Klymkowsky, M. W., & Moody, S. A. (2008). Eya1 and Six1 promote neurogenesis in the cranial placodes in a SoxB1-dependent fashion. *Dev Biol, 320*(1), 199-214. doi:10.1016/j.ydbio.2008.05.523

Sheng, G., & Stern, C. D. (1999). Gata2 and Gata3: novel markers for early embryonic polarity and for non-neural ectoderm in the chick embryo. *Mech Dev, 87*(1-2), 213-216. doi:10.1016/s0925-4773(99)00150-1

Simoes-Costa, M., & Bronner, M. E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development, 142*(2), 242-257. doi:10.1242/dev.105445

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Haring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L., Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., Croci, L., Hsiao, W. Y., Guertin, D. A., Brunet, J. F., Consalez, G. G., Ernfors, P., Fried, K., Kharchenko, P. V., & Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science, 364*(6444). doi:10.1126/science.aas9536

Steventon, B., Araya, C., Linker, C., Kuriyama, S., & Mayor, R. (2009). Differential requirements of BMP and Wnt signalling during gastrulation and neurulation define two steps in neural crest induction. *Development, 136*(5), 771-779. doi:10.1242/dev.029017

Steventon, B., Mayor, R., & Streit, A. (2012). Mutual repression between Gbx2 and Otx2 in sensory placodes reveals a general mechanism for ectodermal patterning. *Dev Biol, 367*(1), 55-65. doi:10.1016/j.ydbio.2012.04.025

Streit, A. (2002). Extensive cell movements accompany formation of the otic placode. *Dev Biol, 249*(2), 237-254. doi:10.1006/dbio.2002.0739

Streit, A., Berliner, A. J., Papanayotou, C., Sirulnik, A., & Stern, C. D. (2000). Initiation of neural induction by FGF signalling before gastrulation. *Nature, 406*(6791), 74-78.

Streit, A., Lee, K. J., Woo, I., Roberts, C., Jessell, T. M., & 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, 125*(3), 507-519.

Streit, A., & Stern, C. D. (1999). Establishment and maintenance of the border of the neural plate in the chick: involvement of FGF and BMP activity. *Mech Dev, 82*(1-2), 51-66.

Stuhlmiller, T. J., & García-Castro, M. I. (2012). FGF/MAPK signaling is required in the gastrula epiblast for avian neural crest induction. *Development, 139*(2), 289-300. doi:10.1242/dev.070276

Stundl, J., Bertucci, P. Y., Lauri, A., Arendt, D., & Bronner, M. E. (2021). Evolution of new cell types at the lateral neural border. *Curr Top Dev Biol, 141*, 173-205. doi:10.1016/bs.ctdb.2020.11.005

Thiery, A., Buzzi, A. L., & 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. doi:10.1016/bs.ctdb.2020.04.002

Trevers, K. E., Lu, H. C., Yang, Y., Thiery, A., Strobl, A. C., Pálinkášová, B., de Oliveira, N. M. M., de Almeida, I. M., Khan, M. A. F., Moncaut, N., Luscombe, N. M., Dale, L., Streit, A., & Stern, C. D. (2021). A gene regulatory network for neural induction. *bioRxiv*, 2021.2004.2016.440164. doi:10.1101/2021.04.16.440164

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., & Streit, A. (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*(2), 355-360. doi:10.1073/pnas.1719674115

Wakamatsu, Y., Endo, Y., Osumi, N., & Weston, J. A. (2004). Multiple roles of Sox2, an HMG-box transcription factor in avian neural crest development. *Dev Dyn, 229*(1), 74-86. doi:10.1002/dvdy.10498

Williams, R. M., Lukoseviciute, M., Sauka-Spengler, T., & Bronner, M. E. (2022). Single-cell atlas of early chick development reveals gradual segregation of neural crest lineage from the neural plate border during neurulation. *Elife, 11*. doi:10.7554/eLife.74464

Wilson, S. I., Graziano, E., Harland, R., Jessell, T. M., & Edlund, T. (2000). An early requirement for FGF signalling in the acquisition of neural cell fate in the chick embryo. *Curr Biol, 10*(8), 421-429.

Yardley, N., & Garcia-Castro, M. I. (2012). FGF signaling transforms non-neural ectoderm into neural crest. *Dev Biol, 372*(2), 166-177. doi:10.1016/j.ydbio.2012.09.006

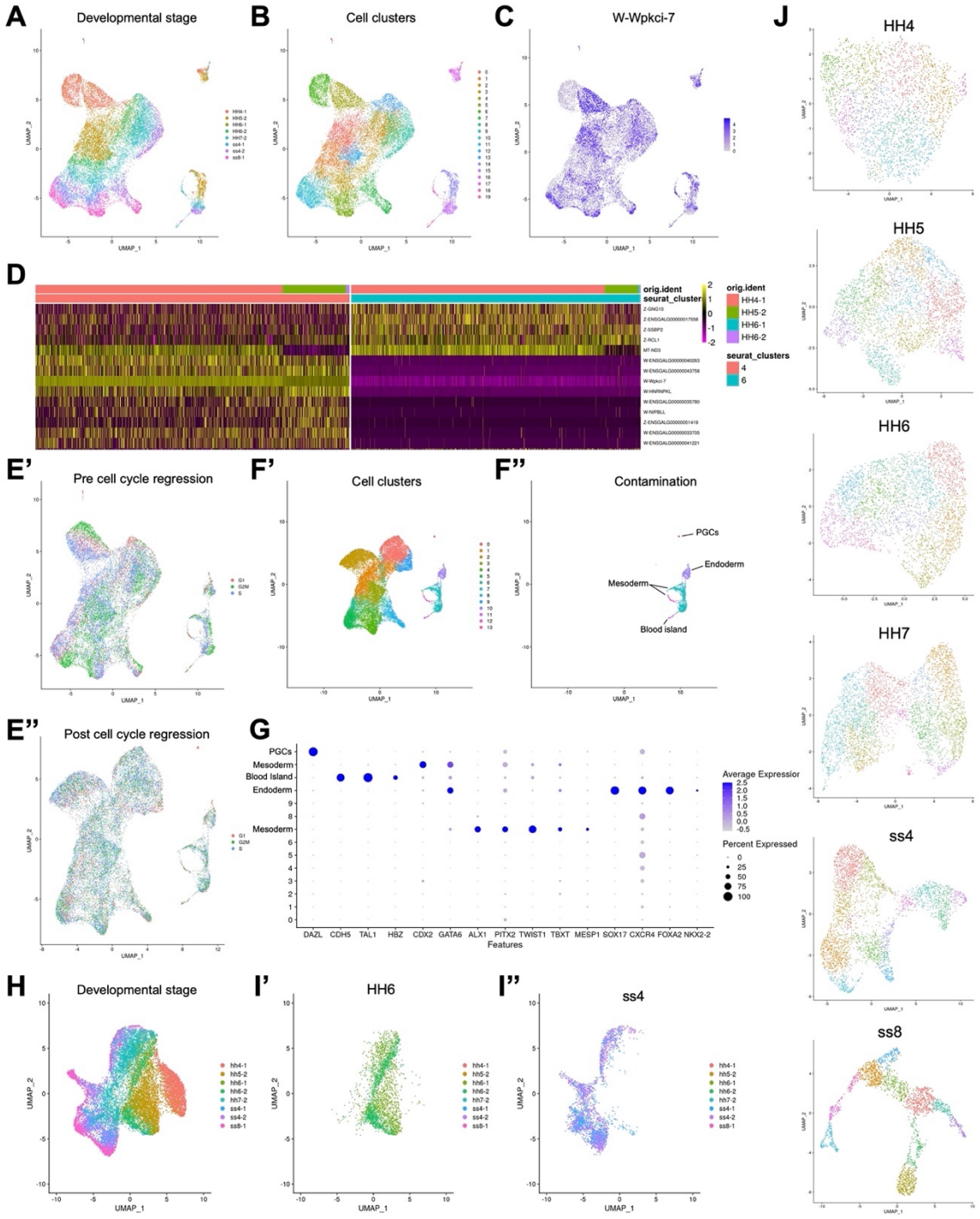


Figure 1-figure supplement 1. (A-B) UMAP of full dataset prior to regressing out confounding factors. Cells are coloured by developmental stage (A) and Seurat cell clusters (B). **(C)** Feature plot displaying W (male) sex-chromosome gene Wpkci-7, illustrating the cell sex effect after clustering. **(D)** Heatmap of selected sex genes across clusters 4 and 6. These clusters contain mostly cells from HH4 and highlight the clear sex effect of W and Z chromosome genes between these clusters. **(E)** UMAP of cells prior to (E') and post (E'') regressing out the cell cycle effect. **(F')** UMAP displaying cell clusters after regressing out sex and cell cycle effects. **(F'')** UMAP highlighting contaminating cell states following cell state classification. These include primordial germ cells (PGCs), mesoderm, endoderm, and blood islands. **(G)**

Dotplot for the expression of genes used for identification of contaminating cell clusters. **(H)** UMAP of all cells after filtering of contaminating populations and regressing out confounding variables. Cells are coloured by developmental stage. **(I)** UMAP plots showing successful integration of two sequencing batches at stages HH6 (I') and ss4 (I''). **(J)** UMAPs displaying cell clusters calculated at each developmental timepoint. Individual stages are subset from the full filtered dataset displayed in (H).

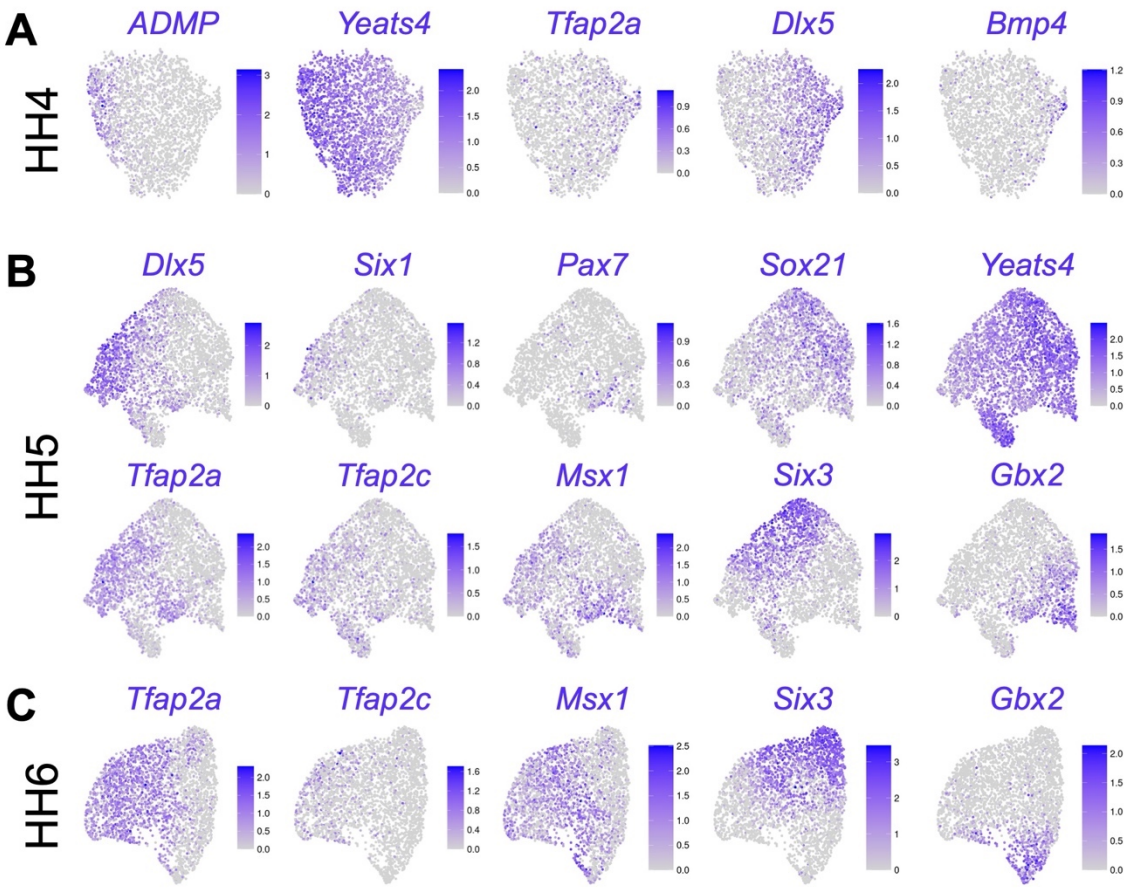


Figure 1-figure supplement 2. Feature plots showing expression of marker genes on UMAPs of developmental stages HH4 **(A)**, HH5 **(B)** and HH6 **(C).**

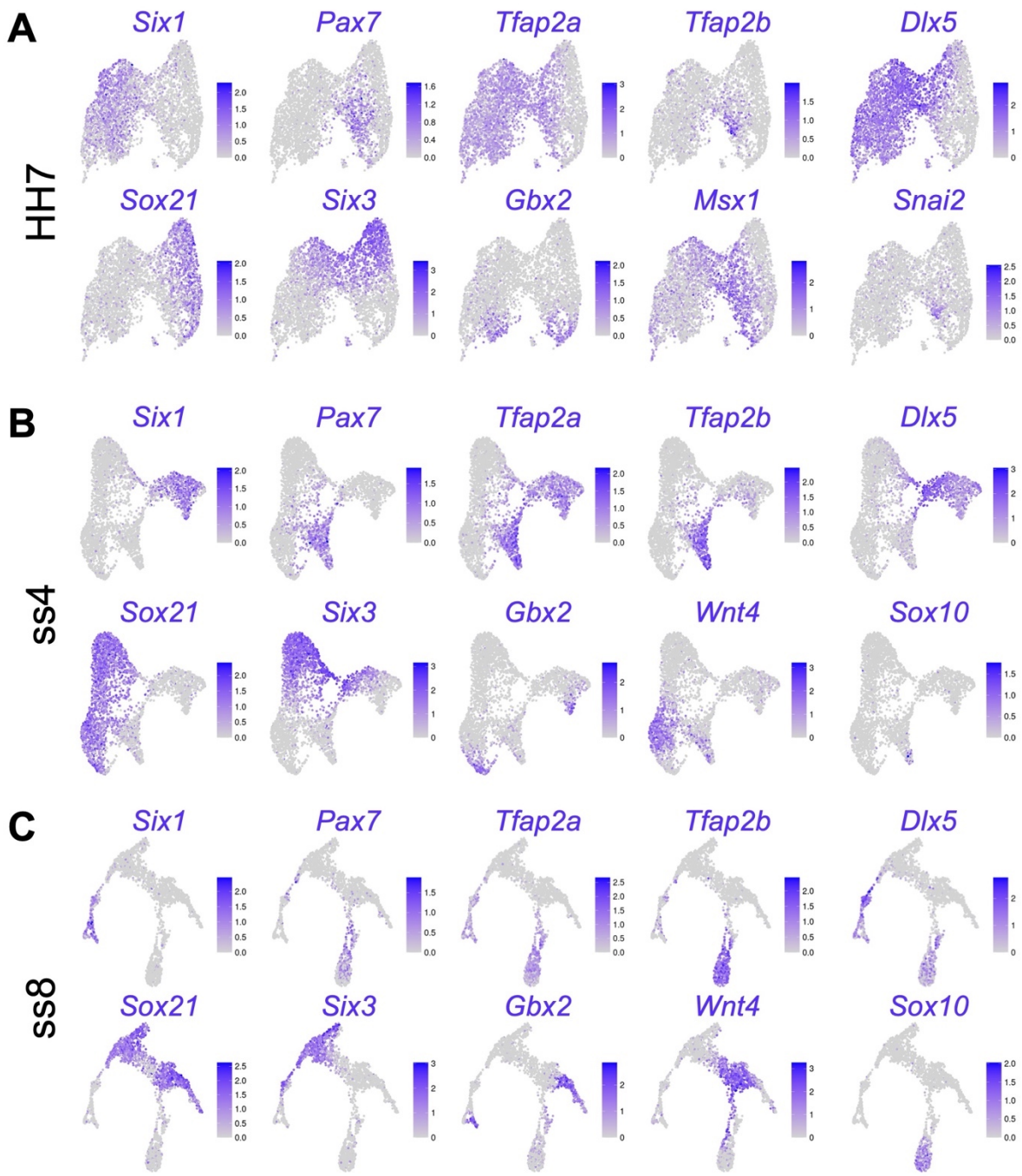


Figure 2-figure supplement 1. Feature plots showing expression of marker genes on UMAPs of developmental stages HH7 **(A)**, ss4 **(B)** and ss8 **(C).**

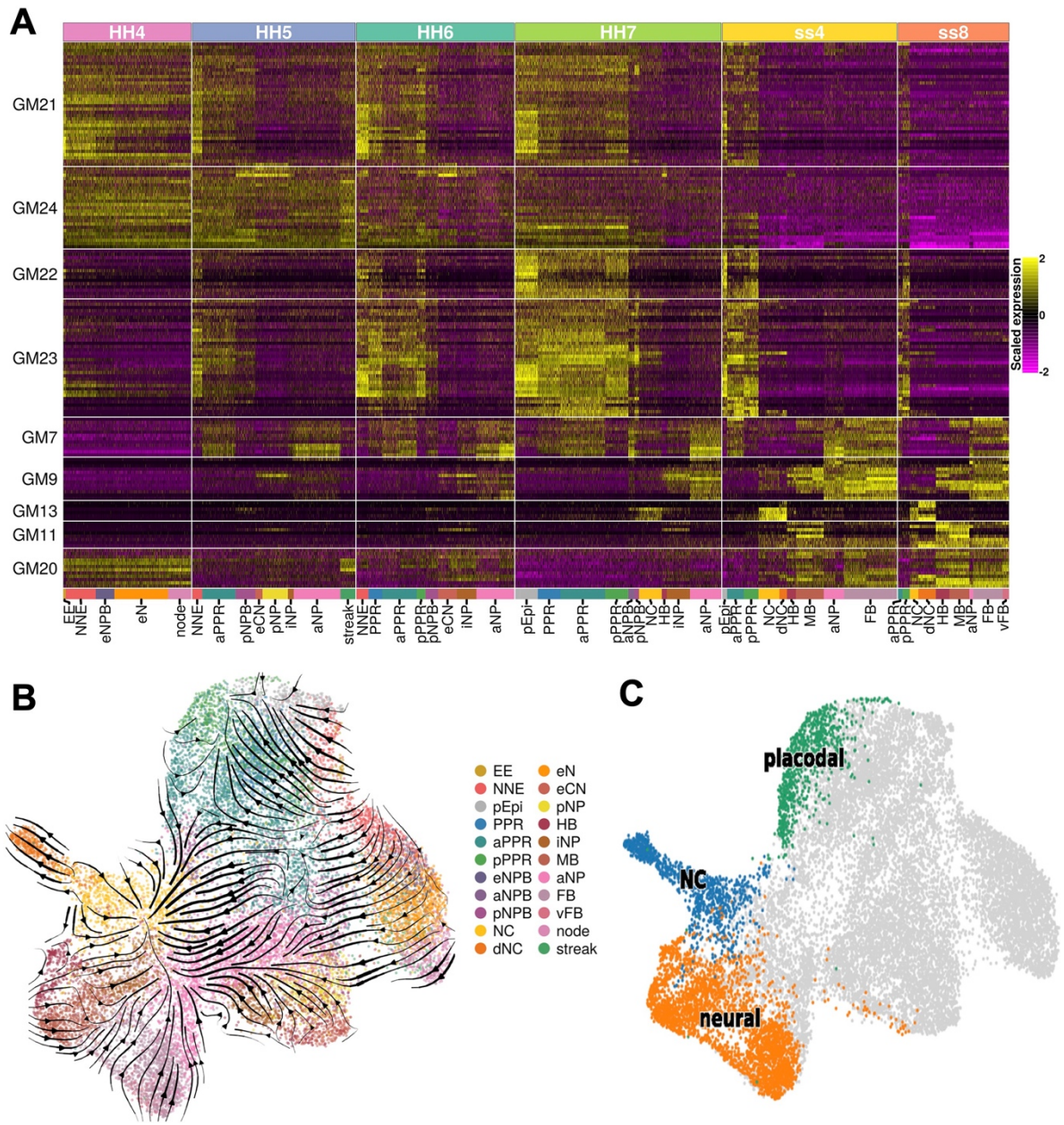


Figure 3-figure supplement 1. (A) Gene modules calculated across the full dataset (full list of gene modules available in Figure 3-Source Data 1). Gene modules were unbiasedly filtered to include those that only display differential expression between cell states and not between sequencing batches and then further filtered to include only those that are differentially expressed between the neural, neural crest and placodal cell states at ss8. 9 gene modules remained after filtering. **(B)** UMAP plot of the full dataset overlaid with RNA velocity vectors depicting the predicted directionality of transcriptional change. Cells coloured by cell state. **(C)** Selected terminal states (placodal, neural crest and neural) used for CellRank lineage inference.

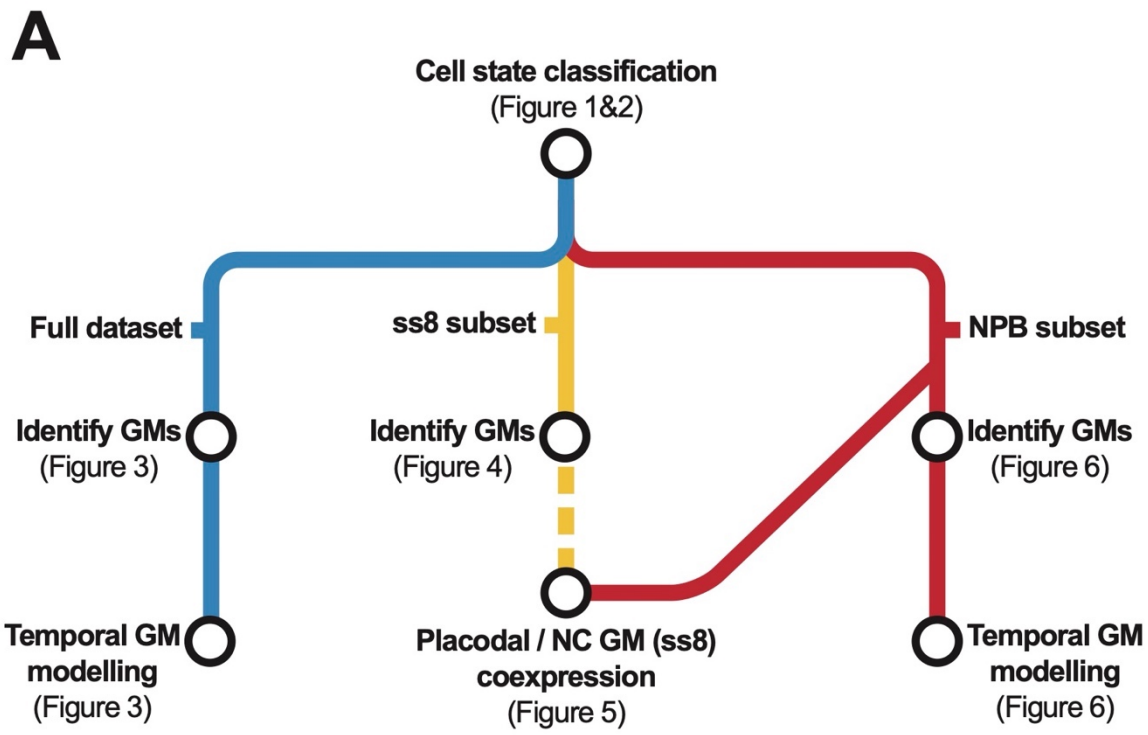


Figure 3-figure supplement 2. (A) Workflow diagram illustrating the different dataset subsets used in each of the main figures.

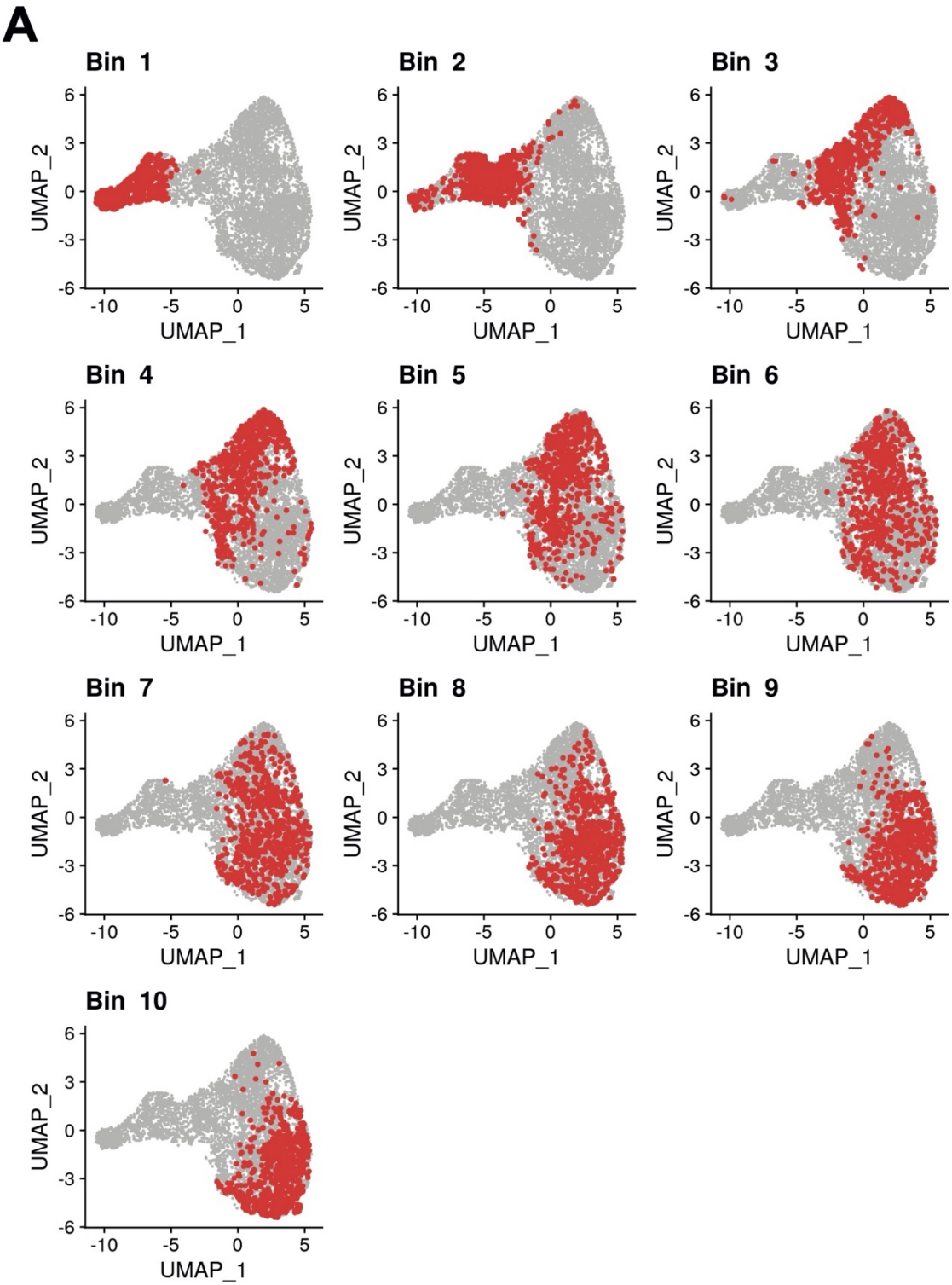


Figure 5-figure supplement 1. (A) UMAPs of the NPB subset. Cells from the NPB subset were ordered by their expression ratio of placodal and neural crest gene modules and then separated into 10 evenly sized bins (Figure 5D). Cells belonging to each of these 10 bins are displayed on the UMAPs.

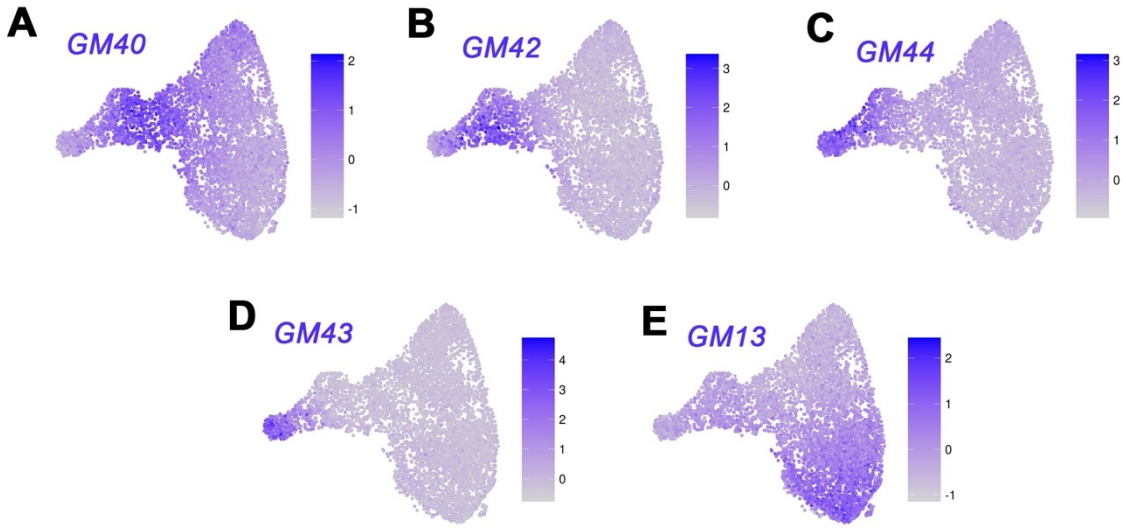


Figure 6-figure supplement 1. (A-E) Feature plots of average gene module expression used for the co-expression analysis in Figure 6I.

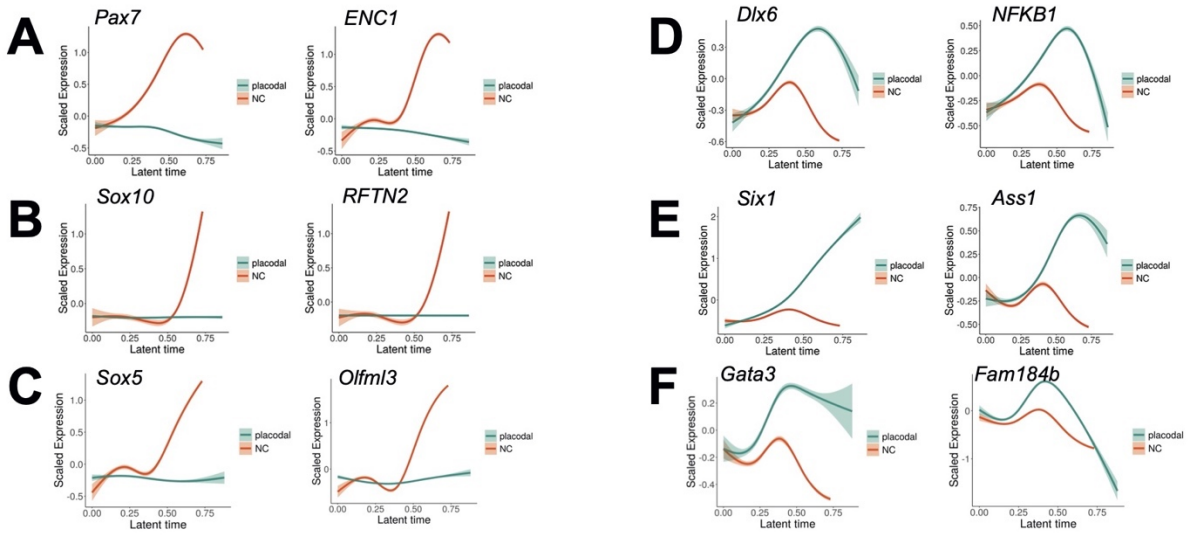


Figure 6-figure supplement 2. (A-F) GAMs are weighted by the absorption probability of each cell towards one of the two fates (placodal or neural crest). Well-characterized genes are shown on the left and novel genes with similar expression dynamics are shown on the right of each panel.