

Neural Crest Migration is Driven by a Few Trailblazer Cells with a Unique Molecular Signature Narrowly Confined to the Invasive Front

by Rebecca McLennan^{1*}, Linus Schumacher^{2*}, Jason A. Morrison¹, Jessica M. Teddy¹, Dennis Ridenour¹, Andrew Box¹, Craig Semerad¹, Hua Li¹, William McDowell¹, David Kay², Philip K. Maini², Ruth E. Baker², and Paul M. Kulesa^{1,3**}

1 Stowers Institute for Medical Research, 1000 E. 50th St, Kansas City, MO, 64110, USA

2 Oxford University, Centre for Mathematical Biology, Mathematical Institute, 24-29 St Giles', Oxford, OX1 3LB, UK

3 Department of Anatomy and Cell Biology, University of Kansas School of Medicine, Kansas City, KS, 66160, USA

***both authors contributed equally to this work, **corresponding author**

SUMMARY

Neural crest cell migration is vitally important to the formation of peripheral tissues during vertebrate development. However, how neural crest cells respond to different microenvironments to maintain persistence of direction and cohesion in multicellular streams remains unclear. To address this, we profiled eight subregions of a typical cranial neural crest cell migratory stream at two distinct phases of migration. Hierarchical clustering showed a significant difference in the expression profile of the lead three subregions compared to newly emerged cells. Multiplexed imaging of mRNA expression using fluorescent hybridization chain reaction (HCR) quantitatively confirmed the expression profiles of lead cells. Computational modeling next predicted a small fraction of lead cells that detect direction information is optimal for successful stream migration. Single cell profiling then revealed a unique molecular signature consistent and stable over time in a subset of lead cells within the most advanced portion of the migratory front, which we term as trailblazers. Model simulations that forced a lead cell behavior in the trailing subpopulation predicted cell bunching near the migratory domain entrance. Mis-expression of the trailblazer molecular signature by perturbation of two upstream transcription factors agreed with the *in silico* prediction and showed alterations to neural crest cell distance migrated and stream shape. These data are the first to characterize the molecular diversity within a neural crest cell migratory stream and offer insights into how molecular patterns are transduced into cell behaviors.

INTRODUCTION

In examples that range from primitive streak formation to mechanosensory organogenesis, several embryonic cell populations undergo persistent, directed migration in coordinated groups (Tarbashevich and Raz, 2010; Voiculescu et al., 2014; Piotrowski and Baker, 2014). When migratory cells fail to reach a target or populate an incorrect location, this often leads to improper cell differentiation or uncontrolled cell proliferation. Thus, studies of embryonic cell migration mechanisms are important for better understanding birth defects and tumor formation.

One of the key features of embryonic cell migration is the persistent, directed movement of cells in multicellular streams. During multicellular streaming, a cell autonomously controls its cytoskeleton but moves with neighbors as a population. A long standing question has been what mechanisms regulate the persistence of direction and cohesion of multicellular streams? Large scale genomic analyses of premigratory or migrating embryonic cells, such as the neural crest, have shed light on the genes expressed in migratory versus non-migratory cells during embryonic development (Gammill and Bronner-Fraser, 2002; Gammill and Bronner-Fraser, 2003; Molyneaux et al., 2004; Adams et al., 2008; Gallardo et al., 2010; Simoes-Costa and Bronner, 2013; Simoes-Costa et al., 2014). However, what remains unclear is how gene expression varies as migrating cells respond to different microenvironments and how this transduces into observed cell behaviors. Therefore, investigative efforts that correlate molecular interrogation with in vivo cell behavior analyses will yield important insights into embryonic cell migration events at the level of both individual cells and the population.

The neural crest is one of the most striking examples of long distance embryonic cell migration that is accessible to manipulation and in vivo observation. Neural crest cells emerge from the dorsal neural tube and are sculpted by intrinsic and extrinsic signals into discrete, multicellular streams throughout the head and trunk (Kulesa and Gammill, 2010). Analysis of cell behaviors from in vivo time-lapse imaging in chick (Kulesa and Fraser, 1998, 2000; Teddy and Kulesa, 2004, Kulesa et al., 2008; McKinney et al., 2011; Ridenour et al., 2014) and in chick and mouse intestine (Young et al., 2004, 2014; Nishiyama et al., 2012) has shown there are regional differences in cell speed, direction, proliferation, calcium activity, and cell morphology depending on cell position within a neural crest migratory stream. We previously showed that a computational model of neural crest cell migration, using the chick cranial neural crest cell behavioral data, predicted that successful cell persistence of direction and stream cohesion would result from the presence of unique lead and trailing subpopulations (McLennan, Dyson et al., 2012). Further, tissue transplantation studies in which lead cells were placed into the trailing stream or vice-versa showed that neural crest cell behaviors and gene expression profiles are not pre-determined but depend on stream position. What remains unclear is how the size and changes and in the molecular profile of the lead neural crest cell subpopulation affects persistence of direction and stream cohesion.

To address these questions, we first isolated neural crest cells from eight discrete subregions of a typical migratory stream in the head using laser capture microdissection (LCM) and analyzed the expression of 77 genes using real time quantitative polymerase chain reaction (RT-qPCR). We examined regional differences

in gene expression and used a newly emerged fluorescent hybridization chain reaction (HCR) strategy for multiplexed imaging of mRNA expression (Choi et al., 2010, 2014). We next isolated and profiled single lead neural crest cells (within the most distal portion of the invasive front) at two distinct phases of migration. We compared single cell gene expression of leaders to cells isolated from the entire stream. We identified a molecular signature unique to these selected lead neural crest cells (which we define as ‘trailblazers’) narrowly confined to the invasive front. We then used gain- and loss-of-function of transcription factors upstream of the trailblazer molecular signature to determine whether genes expressed as part of this signature are critical to neural crest cell migration. These experiments were performed in parallel with computational modeling that simulated our experimental scenarios. Our results highlight the molecular heterogeneity of cells during cranial neural crest cell migration and the requirement for only a small subset of trailblazer cells to ensure persistence of direction and stream maintenance.

MATERIALS AND METHODS

Embryos and cell labeling

Fertilized white leghorn chicken eggs (supplied by Centurion Poultry Inc., Lexington, GA) were incubated at 38 °C in a humidified incubator until the desired Hamburger and Hamilton (HH) stages of development (Hamburger and Hamilton, 1951). Plasmid DNA (2.5-5 ug/ul) or fluorescein tagged morpholinos (0.5 mM) were injected into the lumen of the neural tube and electroporated at HH Stage 9 to affect the entire stream (McLennan and Kulesa 2007). We focused the morpholino knockdowns

to a specific developmental stage of interest with a large number of n values for which we did not observe any obvious off-target effects. For electroporations of the trailing portion of the neural crest stream, embryos were incubated until 10-12 somites at which point plasmid DNA was injected and electroporated with CellTracker CM-Dil (C-7001, Life Technologies). We co-injected H2B-mCherry nuclear label to mark some unperturbed neural crest cells within a morpholino-transfected stream.

Gene profiling

(Eight segment) Tissue was harvested by LCM (Zeiss), pre-amplified using a modified version of the Cells-to-Ct kit (Ambion) and analyzed by microfluidic RT-qPCR on the BioMark HD (Fluidigm) as previously described (Morrison et al., 2012). The stream lateral to rhombomere 4 (r4) (in cryosections) was divided into eight approximately equivalent subregions (the number of subregions selected for our ability to cut reproducibly) and harvested by LCM. Following LCM, RNA from residual cryosections produced RNA integrity numbers (RINs) of 5.8-6.8 on a Bioanalyzer 2100 (Agilent). 77 transcripts were pre-amplified from cDNA using 14 cycles per the Cells-to-Ct protocol (Life Technologies). The majority of genes examined were chosen from neural crest literature. Preamplified cDNAs were diluted with sterile 1X TE and the products analyzed on a Fluidigm BioMark HD at the Children's Hospital Boston IDRC Molecular Genetics Core. No Template negative control reactions failed to produce signal. Median Absolute Deviation (MAD) was used to eliminate outliers, resulting in 3-6 biological replicates per sample. Different outliers were automatically removed when comparing subregions 1-8 than comparing subregions 1-3. A trio of reference genes,

selected from six candidates, was used to calculate normalized relative quantities and differential expression in Biogazelle's qBASEplus software. Hierarchical clustering by Pearson's dissimilarity was performed using Partek's Genomics Suite software package.

(Single cell) For the quartile analysis, four regions of interest (the number of subregions selected for our ability to manually dissect reproducibly) from the cranial r4 migratory stream, electroporated with Gap43-yellow fluorescent protein (YFP), were chilled 0.1% DEPC PBS and mechanically and chemically dissociated. For the trailblazer analysis, the tips of the arches, containing an average of 8 electroporated neural crest cells were manually removed from the rest of the embryo and then mechanically and chemically dissociated. Single, healthy, YFP+, neural crest cells were isolated by fluorescence-activated cell sorting (FACs) into Cells-to-Ct lysis solution with 1:100 DNase I and incubated at room temperature for 20 minutes. Alternatively, single neural crest cells were isolated by LCM from cryosections (McLennan et al., 2012) or from desiccated neural tube cultures (McLennan et al, 2010) on membrane slides (Zeiss). Bioanalyzer 2100 RINs determined from tissue remaining after LCM measured >6.8 for neural tube cultures and >8.9 for cryosectioned tissue. cDNA was synthesized directly from entire lysates using the High Capacity cDNA kit (Life Technologies). An 18-cycle, Cell-to-Ct pre-amplification protocol was employed to selectively amplify 96 transcripts with Taqman Gene Expression Assays (Life Technologies). Pre-amplification products were diluted with sterile 1X TE before being run on Fluidigm's BioMark HD. Raw amplification curves were evaluated manually to ensure correct calling of logarithmic and non-logarithmic curves. Single cell RT-qPCR data were analyzed using

Fluidigm's Singular2 software package in R, including elimination of outliers. Limit of detection (LoD) was set at 28 for the single cell isolation method comparison and 26 for the trailblazer and quartile single cell profiling. For analysis of combined populations of single cells a trio of reference genes, selected from seven candidates, was used to calculate normalized relative quantities and differential expression in Biogazelle's qBASEplus software. Hierarchical clustering and intensity plots were constructed using PARTEK's Genomics Suite.

Fluorescent multiplex in situ hybridization

Transcripts were visualized in whole embryos and tissue sections by HCR. Chick embryos were incubated to specific developmental stages, tissue rapidly collected in chilled 0.1% DEPC PBS and fixed in fresh 4% paraformaldehyde at ambient temperature for 1 hr. Embryos were then dehydrated and rehydrated before HCR was performed according to the manufacturer's instructions (Molecular Instruments, Caltech, USA). Tissue sections were cut on a vibratome (Leica VT1000S). In some cases, immunohistochemistry was performed on the tissue following HCR. All samples were imaged by confocal microscopy (Carl Zeiss, LSM 780) and quantitation of fluorescence signal using polyline kymograph analysis (Jay Unruh, Stowers) in Image J (FIJI).

Cell migration measurements

Segmentation of individual fluorescently labeled neural crest cells was completed in Imaris (Bitplane Inc., Saint Paul, MN) using the spots function. Cells were divided

among equal bins along the migratory route (calculated as 10% of the total stream length in the x- direction defined from the neural tube to the 2nd branchial arch) and the total cell distance migrated and area covered calculated as previously described (McLennan and Kulesa, 2007). Statistical analyses were performed using the Students t-test or 2nd order ANOVA.

Mathematical modeling

To test the logical conclusions of our mechanistic hypotheses, we used a hybrid computational model (McLennan, Dyson et al, 2012), with individually represented cells and a continuous chemoattractant concentration (Fig. 4A-D). In this model, cells undertake a two-dimensional off-lattice random walk on a growing rectangular domain that represents the neural crest cell migratory domain (Fig. 4A-D). New cells enter the domain from one end throughout the simulation to represent neural crest cell emigration from the dorsal neural tube (Fig. 4D). The chemoattractant concentration is modeled by a reaction-diffusion equation, with the cells acting as sinks, representing internalization of chemoattractant. The key model components are illustrated in Fig. 4, Supplementary Fig. 6, and model parameters are listed in Table 1 (further model information in the Supplementary Material).

We extended our original computational model (McLennan, Dyson et al., 2012) to incorporate new experimental findings and performed simulations in parallel with the experiments presented in this paper (Figs. 4, 8). Key changes include:

(1) a wider stream of cells that allows for greater cell numbers and more adequate representation of multicellular stream migration; (2) limiting the sensing accuracy of

cells (for chemotaxis). Based on Berg and Purcell (1977) we derived (order of magnitude) bounds of how small a local chemoattractant gradient can be relative to the bulk concentration before cells cannot sense it; (3) assuming that when cells are in contact, the length of the filopodium giving rise to the contact is not fixed. Instead, a range of intercellular distances is allowed, with a maximum beyond which cells lose contact and cease to communicate directional information. These distances are based on empirical data ([Supplementary Fig. 6](#)) and this change improves stream cohesion and reduces stream break-up in model migration.

RESULTS

Molecular profiling reveals regional diversity in gene expression with a gradual transition in patterning between subregions of the cranial neural crest stream

We previously used computational modeling, cell morphometrics, and molecular profiling to reveal the existence of at least two subpopulations of cells with distinct molecular profiles within a typical cranial neural crest cell migratory stream (McLennan, Dyson et al., 2012). However, it remained unclear whether there were only two distinct gene expression profiles corresponding to lead and trailing cells or whether there was a much richer regional diversity of gene expression patterns that varied from the invasive front to the newly emerged trailing cells.

To address this question, we subdivided the pre-otic, cranial neural crest migratory stream into eight subregions (Fig. 1A-D). Hierarchical clustering of 77 genes revealed distinct differences in gene expression between each of the eight subregions ([Fig. 1E, F](#)). Subregions 1-3, 4-5, and 6-8, clustered together, respectively (Fig. 1F).

When we organized the genes according to their level of linear expression in subregion 1 (invasive front) relative to subregion 8 (newly emerged from the neural tube), we found non-linear transitions of gene expression between the clustered subregions (Fig. 1F). Some of the genes highly expressed within the invasive front (subregion 1) were dramatically diminished in expression towards the proximal trailing subregions (Fig. 1F). Likewise, some genes expressed at low levels within the invasive front (Fig. 1F; bottom half of graph) showed a significant increase in expression when compared throughout the more proximal subregions. The lead three subregions (1-3; corresponding to the invasive front and distal portion of the cranial neural crest cell migratory stream) cluster very closely with one another and are distinct from the newly emerged neural crest cells (subregions 6-8) (Fig. 1G). Of the profiled genes, 44% (34/77) are differentially expressed between these lead three subregions and the more proximal trailing cell subpopulations (subregions 6-8) (Fig. 1G). These data reveal a widespread regional diversity in gene expression profiles based upon cell position within the cranial neural crest cell migratory stream.

Hierarchical clustering shows uniqueness of the distal stream and four characteristic patterns of gene expression

To determine whether the neural crest cells within the most advanced portion of the migratory front display a unique molecular profile, we compared the gene profiles of cells from subregion 1 (lead 12.5% of migratory stream) to those cells immediately proximal in subregions 2-3 (Fig. 1H). The invasive migratory front shows a molecular profile distinct from cells within subregions 2-3 with significant up or down-regulation of

18% (14/77) of the genes analyzed (Fig. 1H). Genes significantly upregulated in subregion 1 compared to subregions 2 and 3 included heart and neural crest derivatives expressed 2 (HAND2), aquaporin 1 (*AQP1*), BMP and activin membrane-bound inhibitor homology (*BAMBI*), glypican 3 (*GPC3*) and matrix metalloproteinase 2 (*MMP2*) (Fig. 1H). When we compared the molecular profile of lead (subregion 1) neural crest cells to the rest of the migratory stream (subregions 2-8), we found similar genes were enriched in the invasive front (*HAND2*, *GPC3*, *MMP2*) (Fig. 1I). Thus, there is a unique molecular profile of the most invasive migratory front.

To determine whether the neural crest molecular profiles along the migratory stream could be categorized, we examined the profiles of individual genes along the migratory stream (Fig. 2). First, we found a pattern where genes were more highly expressed in the invasive front and expression decreased proximally to the newly emerged neural crest cells (Fig. 2A). Second, we observed a pattern that peaked in expression in the subpopulation of migrating cells within the middle portion of the migratory stream (Fig. 2B). Third, there are genes more highly expressed in the newly emerged neural crest cells in a pattern that diminishes towards the migratory front (Fig. 2C). Lastly, some genes show a uniform expression pattern across all subregions of the cranial neural crest cell migratory stream (Fig. 2D). Thus, the regional differences in gene expression in the cranial neural crest cell migratory stream vary, displaying four distinct patterns.

Multiplexed imaging of mRNA expression using fluorescent HCR confirms the regional patterns of gene expression within the cranial neural crest cell migratory stream

To confirm regional differences in gene expression within the neural crest cell migratory stream, we applied a newly developed technology known as fluorescent HCR (Choi et al., 2010, 2014). HCR takes advantage of initiator probes and hairpins to detect and amplify mRNA signals within cells and provides a fluorescent readout of mRNA expression (Choi et al., 2010, 2014). This allows more accurate quantitative analysis and multiplexed imaging of mRNA expression to better resolve gene expression patterns and borders than traditional in-situ hybridization. We applied HCR to simultaneously visualize the expression profiles of two genes highly expressed by cells within the most invasive front (*HAND2* and *BAMBI*) and combined HCR with HNK-1 immunolabeling (to distinctly detect neural crest cells) (Fig. 3A,B). Visual analysis revealed the patterns of *BAMBI* and *HAND2* expression that confirmed restricted expression in the lead neural crest cells and mid-stream subregions at HH Stage 15 as indicated by HCR (Fig. 3A, C). Further, quantitative analysis of the fluorescent HCR signals using a polyline kymograph method confirmed the patterns of expression of *BAMBI* and *HAND2* (Fig. 3D). Thus, multiplexed imaging of mRNA expression using fluorescent HCR allowed us to quantitatively confirm regional differences in gene expression within the cranial neural crest cell migratory stream.

Computational modeling predicts that a small fraction of lead cells is optimal for successful stream migration

We use our extended computational framework to test aspects of our hypothesis that are difficult to probe in vivo. Specifically, we asked whether the number of lead cells that can detect spatial gradients in chemoattractant is a critical factor for the success of neural crest cell invasion. Should this be the case, then the gene expression patterns detailed in vivo (Figs. 1,2) would be consistent with this constraint. To test the effect of varying the number of leaders, we first restricted the model simulations to only include non-plastic cell behaviors, such that individual cells that start migration as leaders (or followers) could not switch from being a leader to a follower (or vice versa). To change the fraction of all migrating cells that are leaders, we varied the time, t_{LF} at which new cells entering the domain are prescribed followers instead of leaders (Fig. 4I): that all cells that entered the migratory domain up to time t_{LF} were specified as leaders, and cells that entered after that were specified as followers.

Our simulations reveal that the median distance migrated and the stream density both increase with decreasing leader fraction (Fig. 4I; compare Movie S1 with low (0.051) vs high (0.59) mean leader fraction). While the furthest distance migrated does not change noticeably, it is the movement of cells away from the entrance to the migratory domain (corresponding to near the dorsal neural tube) that proves critical for the successful migration pattern (Fig. 4I,J). This prevents jamming near the domain entrance and enables a higher number of cells to distribute more evenly along the migratory route (Fig.4I, J). This trend cannot be extrapolated to zero leaders, in which case the followers move in a random, undirected manner and stay close to the entrance of the domain (Movie S2). Thus, our computational model predicts that a small number

of lead cells can efficiently guide the migration of the entire neural crest cell migratory stream.

Single cell analysis of the invasive front identifies a subpopulation of neural crest cells that we define as trailblazer cells with a stable and consistent set of highly expressed genes

To test our computational model prediction of a small leader fraction within the invasive front, we performed single cell analysis and looked for evidence of a stable and consistent set of highly expressed genes within lead neural crest cells. To do this, we isolated an average of 8 lead neural crest cells per embryo (by manually removing tissue containing the first few fluorescently labeled neural crest cells from the rest of the tissue, cell dissociating that tissue and then sorting via FACS) directly from the most invasive front as cells migrated towards (HH Stage 13), and invaded (HH Stage 15) the branchial arch region (Fig. 5A). We determined FACS isolation to be the most efficient method (as compared to LCM both in vitro and in vivo) for analyzing single neural crest cells while maintaining the native molecular profile (Supplementary Fig. 2). When we profiled 96 genes (composed of the 77 genes previously used for profiling plus additional genes of interest), we discovered a stable and consistent set of genes, described below, expressed by a small fraction of the lead neural crest cells that we define as ‘trailblazers’ (Supplementary Table 2).

Trailblazers have a high degree of gene expression homogeneity during both phases of migration examined, as can be seen in the profiles of individual cells (Fig.

5A). Principle component analysis (PCA) plots of the genes during each phase of migration analyzed show that over 70% of the genes are stably expressed with little variation (Fig. 5B, blue squares; Supplementary Table 2). Violin plots confirm similar levels of expression for a range of genes, including *BAMBI*, *NOTCH1* and *CXCR1* (Fig. 5C). When we focused our attention on the highest expressed genes with RT-qPCR Ct values of less than 22, we found that 98% (61/62) of the most highly expressed genes within the invasive migratory front at HH Stage 15 are also highly expressed at HH Stage 13 (Supplementary Table 2). Many of these genes are consistently expressed by a large percentage of the profiled trailblazers at both phases of migration, including *BAMBI*, *CXCR1*, *NOTCH1*, plakophilin 2 (*PKP2*) and transcription factor AP-2 alpha (*TFAP2A*) (Fig. 5B, C; Supplementary Table 2).

Single cell analysis also shows genes whose expression patterns were dramatically different in the trailblazers when comparing the phases of migration analyzed (Fig. 5B, C). For example, *HAND2* is expressed in a small number of lead neural crest cells (7 %) during the first phase of migration (Fig. 5C, Supplementary Table 2). This dramatically increases during the second phase to 59% of the lead cells (Fig. 5C, Supplementary Table 2). In contrast, the percentage of neural crest cells highly expressing *Sox10* dramatically decreases (from 69% to 21%) over time (Fig. 5C, Supplementary Table 2). *BAMBI* and *HAND2* expression was confirmed at different stages of migration with HCR (Supplementary Fig. 1C). Migratory and pre-migratory neural crest cells were visualized with either FoxD3 or HNK-1 expression. As the first neural crest cells begin to delaminate from the dorsal neural tube, they express *BAMBI* (Supplementary Fig. 1C (box); HH Stage 9). There are a small number of cells at the

lead of the neural crest migratory stream that express high levels of BAMBI at all phases of migration examined (Supplementary Fig. 1C (box); Figure 3). HAND2 was high in forming cardiac regions, but was not expressed in the lead neural crest until HH Stage 15 (Supplementary Fig. 1C (box), Fig. 3), coinciding to the final phase of migration. Therefore, we restricted our focus to genes that are expressed highly by at least 50% of the cells at both time points (Supplementary Table 2, bolded genes). Together, these data confirm a stable and consistent set of genes expressed within the trailblazers.

Trailblazers have a unique molecular signature and are narrowly confined within the invasive migratory front

To determine the molecular signature unique to the trailblazers, we compared the set of genes stably and consistently expressed by the trailblazers to those expressed by neural crest cells within the entire stream at HH Stage 15 (manually dissecting into four quartiles) (Fig. 6A). We measured the similarity of molecular profiles across 318 single neural crest cells by PCA, violin plots, hierarchical clustering, pairwise correlation and Pearson's correlation (Fig. 6BD, Supplementary Fig. 3, 4). Our analysis reveals four main results. First, the vast majority of cells from the quartiles have a poor correlation with the molecular profile of the trailblazers (Fig. 6B). At the single cell level, PCA shows incomplete overlap between genes expressed by trailblazers and quartile 1 (Fig. 6B). Second, neural crest cells within each quartile have high correlations with one another (Fig. 6B). Third, the gene expression profile of the trailblazers is distinct from all quartiles, but most similar to quartile 1 (Fig. 6B). Fourth, the average expression of

neural crest cells in the quartiles cluster according to their position within the migratory stream as well as within the local microenvironment depending on the type of clustering method (Fig. 6D; Euclidean, Pearson). Violin plots reveal examples of individual genes that are expressed at higher (*BAMBI*, *CXCR1*, *PKP2*, *HAND2*) or lower (*Sox10*, integrin alpha 3 (*ITGA3*)) levels by trailblazers than cells located more proximal within the migratory stream (Fig. 6C). Measurements of the *HAND2* and *BAMBI* HCR expression at higher resolution within the four quartiles confirmed the regional differences in expression (Supplementary Fig. 1A, B). Thus, the unique molecular signature associated with trailblazer cells is distinct from cells in the entire first quartile and is not shared by other migrating neural crest cells within the rest of the stream.

To refine the unique molecular signature of the trailblazers, we examined the genes that were differentially expressed between trailblazers and quartile 1 at HH Stage 15 (Supplementary Table 3, bolded genes). This resulted in a list of 17 genes highly expressed in trailblazers and differentially expressed between the trailblazers and migrating neural crest cells in quartile 1. This list of genes did not take into account comparison of the trailblazers with the remainder of the stream. Therefore, a gene expressed highly by cells in quartile 3 and trailblazers would remain in the trailblazer molecular signature. To remove any such genes, we compared the invasive front profile to single cell profiles from cells isolated throughout the stream (quartiles 2-4) and found 94% (16/17) of the genes in agreement (Supplementary Table 3, bolded genes). Thus, these 16 genes are representative of the unique molecular signature of the trailblazer cells.

The transcription factors *HAND2* and *TFAP2A* are identified as likely to influence the unique molecular signature associated with trailblazers and the neural crest cell migration pattern

To test the function of the genes enriched in the invasive front during neural crest cell migration, we examined the gain and loss of function of *TFAP2A* and *HAND2*. We selected *TFAP2A* and *HAND2* based on the fact that they are transcription factors and their enrichment within cells in the lead three subregions (Fig. 1G). *HAND2* and *TFAP2A* are also expressed at high levels by cells within the invasive front at both phases of neural crest cell migration analyzed (Supplementary Table 2). Lastly, all of the genes in the unique molecular signature of the invasive front are regulated by either *HAND2*, *TFAP2A* or by both (Fig. 9).

Disruption of *HAND2* and *TFAP2A* function alters the unique molecular signature associated with the trailblazers

To determine the changes in the unique molecular signature associated with the trailblazers (Fig. 9) to the loss of *HAND2* or *TFAP2A* function, we profiled migrating neural crest cells 24 hours after transfection with *HAND2* or *TFAP2A* morpholinos (HH Stage 15). We discovered that when *HAND2* is knocked down, several trailblazer signature genes are significantly up-regulated including *AQP1*, *CDH11*, *CDH7*, *CXCR4* and *EPHB1* (Supplementary Table 4, bolded genes). We also learned that *TFAP2A* is down-regulated in *HAND2* morpholino transfected embryos, indicating crosstalk between the two transcription factors (Supplementary Table 4). Loss of *TFAP2A*

function results in the up-regulation of integrin beta 5 (*ITGB5*) and *NEDD9*, and down-regulation of *CXCR4* and *EphB1* (Supplementary Table 4, **bolded genes**). Loss of either *HAND2* or *TFAP2A* results in up-regulation of *HAND2* expression, suggesting either negative regulation or activation of a compensatory pathway. From this, we conclude that loss of either *HAND2* or *TFAP2A* function within migrating neural crest cells influences the unique molecular signature associated with the trailblazer neural crest cells.

Knockdown of *HAND2* or *TFAP2A* expression results in alterations to the neural crest migration pattern

We hypothesize that genes expressed by trailblazers are critical to cell direction persistence and stream migration. To test this hypothesis, we knocked down *TFAP2A* and *HAND2* function in cranial neural crest cells using morpholinos. We find that loss of *HAND2* function results in a significant reduction in the area invaded by treated versus control neural crest cells (Fig. 7A, F,G). Upon close examination, the fluorescence associated with individual neural crest cells within *HAND2*-morpholino embryos was more punctate than in control morpholinos (Fig. 7A). This phenotype does not correlate with increased cell death (Supplementary Fig. 5), and hence we conclude no significant changes in neural crest migration when transfected with *HAND2*-morpholino as some cells are able to migrate into the distal portion of the branchial arches (Fig. 7A-E).

Loss of *TFAP2A* function results in dramatic alterations to the cranial neural crest cell migratory pattern (Fig. 7A). First, the distribution of neural crest cells along the

migratory pathway is perturbed when *TFAP2A* is knocked down at both phases of migration due to a statistically significant drop in cell number in the lead subregion of the migratory stream (Fig. 7B-C). During the first phase of migration (HH Stage 13), the distance migrated and area invaded by *TFAP2A* morpholino transfected neural crest cells are not statistically different to control embryos (Fig. 7D, F). However, during the second phase of migration (HH Stage 15), neural crest cells transfected with the *TFAP2A* morpholino stopped and failed to migrate the entire distance of the migratory route (Fig. 7E,G). From these results, we conclude *TFAP2A* is necessary for neural crest cell migration into the branchial arches.

Computational modeling that simulates overexpression of genes associated with the trailblazers within trailing cells predicts alterations to the neural crest migration pattern

To model the overexpression of a gene associated with the unique molecular signature of the trailblazers within the trailing subpopulation, we modified the model parameters to convert 50% of trailing cells at random into leaders. This mimics, in silico, the transfection of 50% of the neural crest cells within the trailing subpopulation with overexpression of the transcription factors of the unique molecular signature associated with the trailblazers. When we forced trailing cells in silico to display a lead cell phenotype, we found that cells remain near the entrance of the migratory domain (Fig. 8A, B). This migratory pattern is similar to the model simulations scenario that introduced higher lead cell fractions (Fig. 4I,J). Thus, forcing a lead cell behavior within

the trailing subpopulation in silico predicts disruption to neural crest cell migration that would be visible as cell bunching near the dorsal neural tube exit.

Overexpression of *HAND2* and *TFAP2A* alters the neural crest migration pattern in a manner consistent with computational model predictions

To experimentally test whether gain-of-function of *TFAP2A* or *HAND2* in the trailing subpopulation of neural crest cells would affect the migration pattern, we selectively overexpressed these genes in the trailing portion of the migratory stream only and measured changes to the migration pattern at HH Stage 15 (Fig. 8C, D). Overexpression of *TFAP2A* in the trailing portion only did not change the distribution of the transfected neural crest cells along the migratory pathway, compared with control EGFP transfected embryos (Fig. 8C-E). However, when *HAND2* is overexpressed in trailing neural crest cells, there are fewer migrating cells but the cells distribute more evenly along the migratory pathway when compared to the distribution of control EGFP transfected neural crest (Fig. 8C-E). This result agreed with our model simulations (Fig. 8A, B) and previous tissue transplantation experiments that showed some lead cells placed into the trailing stream migrated to distribute along the migratory pathway (McLennan, Dyson, et al., 2012). Thus, forcing the expression of HAND2 or a lead cell phenotype within trailing neural crest cells resulted in fewer cells along the migratory route.

DISCUSSION

In this study, we examined the molecular heterogeneity within a typical neural crest cell migratory stream, and the role of genes associated with lead cells to ensure persistence of direction and stream cohesion. Using novel methods to isolate and profile single and small numbers of cells, we analyzed and compared gene expression patterns within distinct subregions of a typical cranial neural crest cell migratory stream during migration towards and colonization of the branchial arches. We identified a consistent and stable molecular signature unique to the cells within the most distal portion of the migratory front, which we termed as trailblazers. We performed gain- and loss-of-function experiments to test whether the regional gene expression differences we identified are necessary for neural crest cell persistence of direction and stream maintenance. In parallel, we used a hybrid computational model to simulate and predict experimental outcomes that tested the size of the lead neural crest cell subpopulation and role of two transcription factors upstream of the unique molecular signature of leaders.

Our work aimed to better understand how embryonic migratory cells respond to microenvironmental cues and whether knowledge of those cues might help predict invasive behaviors. We discovered that extrinsic signals within the various microenvironments through which migrating cranial neural crest cells travel as well as neural crest cell-cell interactions appear to regulate the gene expression profile of cells. Cranial neural crest cells have distinct gene expression profiles that cluster in similarity depending on stream position (Fig. 1,2). These measurements demonstrate that neural crest cells are dynamically responsive to different microenvironmental cues and suggest

that an understanding of these signals may provide insight into the invasive abilities of multicellular stream migration.

One key question was concerned with determining the size of the leaders required to guide neural crest cell persistence of direction and stream maintenance over long distances. Our previous computational model predicted at least two separate cell subpopulations (leaders/trailers), which was confirmed by RT-qPCR analysis that divided the stream into a 30/70 percentage split (McLennan, Dyson et al., 2012). What was unclear and difficult to manipulate experimentally was whether a change in the number of leaders would affect stream dynamics. This is where our extended computational model proved very useful. In fact, we were able to show that the furthest distance migrated was insensitive to the number of lead cells (Fig. 4). Even a few lead cells could migrate as far as the entire multicellular stream (Fig. 4). Not only that, but higher numbers of leader cells were less efficient at guiding the entire stream. This suggests that only a few lead neural crest cells with guidance information are required for persistence of direction and stream maintenance.

Further analysis of lead neural crest cells provided insights into the molecular characteristics of the invasive front. Our single cell analysis led to the identification of a unique molecular signature associated with a few lead cells we termed, trailblazers (Fig. 5, 6). Trailblazers are narrowly confined to the most invasive front. The molecular signature of the trailblazers is distinct from the molecular profiles of other migrating cells within the stream (Fig. 9). This unique molecular signature associated with trailblazers included 16/96 genes involved in cell-cell adhesion (*CDH7*, *CDH11*, *CTNNB1*, and *PKP2*), cell-extracellular matrix (ECM) adhesion (*ITGB5*), differentiation (*NEDD9*,

NOTCH1, *SNAI2*), guidance (*CCR9*, *CXCR1*, *CXCR4*, *CXCR7*, *EphB1*, *EphB3*), proliferation (*BAMBI*) as well as cell shape dynamics (*AQP1*). It is important to note that the other 80/96 genes did show changes in expression over time in the trailblazer cells, suggesting a majority of the molecular profile of trailblazers is influenced by the microenvironment. Thus, computational model simulations that predicted only a few cells are required to direct stream migration in the presence of a chemoattractant on a growing domain led to the discovery of a small subpopulation of lead cells narrowly confined to the invasive front with a consistent and stable molecular signature.

Based on our identification of a unique molecular signature associated with trailblazer neural crest cells, we next asked whether trailblazer gene function is important for neural crest cell persistence of direction and stream maintenance. Since *HAND2* expression was high in lead cranial neural crest cells, we expected that knockdown of *HAND2* would affect the migration pattern. However, neural crest cells with *HAND2* knockdown did not show a migratory defect (Fig. 7). This may be explained though by the differences in *HAND2* expression levels at the two distinct phases of migration analyzed. *HAND2* expression is high at HH Stage 13 but only in 8% of the trailblazer cells. By HH Stage 15, *HAND2* is expressed highly in 59% of trailblazers. Thus, we conclude that knockdown of *HAND2* does not affect early phases of cranial neural crest cell migration, but may play a role in later neural crest cell patterning as has been previously shown (Thomas et al., 1998; Howard et al., 1999; Srivastava et al., 1997; Miller et al., 2003; Hendershot et al., 2007; D'Autreaux et al., 2007).

When the molecular signature of the trailblazers is perturbed via *TFAP2A* morpholino, both their migration and the migration of the following neural crest cells is affected. *TFAP2A* has previously been shown to be expressed by premigratory and migratory cranial neural crest cells and be important in neural crest induction, proliferation and differentiation in mice, zebrafish and *Xenopus*, as well as avian facial tissue growth (Mitchell et al., 1991; Chazaud et al., 1996; Schorle et al., 1996; Shen et al., 1997; Pfisterer et al., 2002; Brewer et al., 2002; Luo et al., 2003; Knight et al., 2003; Knight et al., 2004; Barrallo-Gimeno et al., 2004; Li and Cornell, 2007; Wang et al., 2011). Knockdown of *TFAP2A* resulted in the failure of neural crest cells to migrate completely to the branchial arches (Fig. 7). We have previously shown that when leaders are prevented from migrating into the target site by a physical barrier, trailing neural crest cells sense the paused leaders, reroute around the barrier and become the new leaders (Kulesa et al., 2005). Here, because neural crest cells throughout the stream were transfected with the *TFAP2A* morpholino, no cells were able to take on the role of the trailblazers and migration was hindered (Fig. 7). These data support our conclusion that only a small number of trailblazers are necessary to guide the neural crest stream over long distances.

Computational model simulations that tested the gain-of-function of trailblazer genes within the trailing subpopulation predicted alterations to the neural crest migration pattern (Fig. 8). When similar experimental perturbations were performed in ovo, we observed alterations to the neural crest migration pattern (Fig. 8). That is, when *HAND2* was overexpressed in the trailing neural crest cells, fewer cells were observed along the migratory pathway (Fig. 8). We initially interpreted this result as a possible neural crest

cell delamination defect. However, after comparing the phenotype to computational model simulations (the first 50um of the computer model migratory domain corresponds to neural crest cell migration from the dorsal midline into the paraxial mesoderm), it is more likely that trailing neural crest cells that overexpress *HAND2* properly delaminate but then fail to persist in directed migration. Therefore, we hypothesize that *HAND2* expression in vivo contributes to the trailblazer phenotype. This is supported by the computational model simulations that show that leaders within the trailing portion of the stream fail to migrate due to a lack of chemoattractant to follow. Trailing neural crest cells that look for and follow the stalled leaders also become stalled as a result. Although, we do not have direct in vivo evidence that lead cells detect spatial gradients of chemoattractants differently from trailers, our results clearly show that association of lead cell behaviors with genes expressed at the invasive front is valid. In contrast, when *TFAP2A* was overexpressed in the trailing portion of the migratory stream, no obvious defects were observed (Fig. 8). Since *TFAP2A* was more broadly expressed in the lead subregions of the neural crest cell migratory stream and not restricted to the invasive front (Fig. 1H), this result was not surprising.

Theoretical testing of the effect of the number of lead cells to drive multicellular stream migration required that we restrict our in silico experiments to the case of non-plastic lead and trailing cell behaviors. We did this even though our previous tissue transplantation experiments showed chick cranial neural crest cells alter their gene expression profile depending on cell position within a stream. Here, we disabled phenotype switching in our computational model simulations, for otherwise we would have only been able to change the leader fraction transiently, before phenotype

switching restored it towards the unperturbed case. Alternatively, we could have set up the computational model simulations with varying initial lead cell fractions and, given appropriate microenvironmental parameters, observe what leader fraction emerges naturally from phenotype switching. However, this would require at least a phenomenologically correct implementation of switching, which in turn has to be verified experimentally. This is outside the scope of the current manuscript, but will be addressed in future work.

In summary, we show that the embryonic neural crest microenvironment regulates the gene expression profile and pattern of cranial neural crest cell migration in a manner that is dependent on cell position and phase of migration. This regulation was identified by the presence of a unique molecular signature associated with trailblazer neural crest cells that are narrowly confined to the invasive front. These data support the hypothesis that a few lead neural crest cells interpret complex microenvironmental signals differently than other migrating neural crest cells within the multicellular stream. To test our hypothesis and importance of the unique molecular signature, we showed that mis-expression of transcription factors *TFAP2A* and *HAND2* results in significant alterations to neural crest cell persistence of direction and stream maintenance. We speculate, based on experiment and computational model simulations, that few trailblazers drive migration, and perturbations in the identity and function of trailblazers led to migration defects. When a gene associated with the trailblazer signature is over-expressed (experiment) or the trailblazer cell behavior is forced within the trailing subpopulation (simulation), cells are stranded near the migratory domain entrance. We postulate this is due to a lack of guidance information. These findings were made

possible by studying neural crest cellular and molecular dynamics within the embryonic microenvironment using a closely integrated experimental and theoretical approach.

ACKNOWLEDGEMENTS

PMK would like to thank the Stowers Institute for Medical Research for their kind generosity. LJS would like to thank Louise Dyson for helpful discussions and access to the original code and gratefully acknowledges the U.K.'s Engineering and Physical Sciences Research Council (EPSRC) for funding through a studentship at the Life Science Interface programme of the University of Oxford's Doctoral Training Centre. REM and JAM would like to thank the Histology Core at the Stowers Institute for Medical Research for assistance with cryosectioning.

FIGURE LEGENDS

Figure 1. Distinctive regional expression profiles exist within the cranial neural crest (NC) migratory stream. (A-B) Stereotypical cranial NC migratory streams at HH St15. (C) The NC stream at HH St15 divided into eight subregions by laser capture microdissection. (D-E) Subregions of the HH15 cranial NC stream were analyzed for the expression of 77 genes. (F) Hierarchical clustering of the eight subregions from the HH Stage 15 cranial NC stream by Pearson dissimilarity based upon the 77-gene profile. All samples shown are relative to subregion 8 with genes ordered by their level of linear expression in subregion 1. (G) Heat map of gene expression differences between

leading subregions 1-3 relative to trailing subregions 6-8. (H) Heat map gene expression differences between leading subregion 1 relative to subregions 2-3. (I) Heat map gene expression differences between leading subregion 1 relative to the rest of the stream. e, eye; nc, neural crest; nt, neural tube; r, rhombomere; ba, branchial arch; n, notochord.

Figure 2. Individual genes have regional expression differences within the cranial NC migratory stream. (A-D) Examples of individual genes with expression increased at the invasive front (A), increase in the middle (B), increased in the trailing portion (C) or uniform (D). All graphs are linear expression with error bars depict the standard error. Each subregion is represented by 3-6 biological replicates. (A') Key indicating the eight subregions of the r4 migratory NC stream. (E) Line graph of normalized expression along the r4 migratory NC stream for genes in (A-C).

Figure 3. Fluorescent hybridized chain reaction (HCR) and immunolabeling confirm regional differences in gene expression within the cranial neural crest cell migratory stream. (A) Whole embryo staining and expression pattern of HAND2 and BAMBI HCR in the head at HH Stage 15. The arrowheads point to the first and second branchial arches. (B) Schematic of HCR and immunolabeling application and imaging. (C) Transverse view of HNK-1 antibody staining, and BAMBI and HAND2 HCR at the pre-otic level of a HH Stage 15 embryo. (D) Polyline kymograph analysis (in a 20-um wide band) of the BAMBI and HAND2 fluorescent HCR signal intensity along the neural crest cell migratory pathway (n=5), an example of the pathway analyzed is in (D'). The scale bars are all 100um. nt, neural tube; ov, otic vesicle; ba, branchial arch.

Figure 4. Computational experiment suggests multicellular stream migration can be more successful with fewer leader cells. The basic components of our computational framework are: (A) Cell-induced gradient - Cells internalize chemoattractant near them. (B) Chemotaxis - Leader cells (orange) randomly sample directions and move up the chemoattractant gradient. (C) Leader-follower behavior - Follower cells (purple) randomly search for other cells and, when in contact with another cell, adopt their direction of movement. (D) Migratory domain - Cells are inserted into the start of the domain ($x = 0\mu\text{m}$) throughout the simulation, but only if there is sufficient space, so that cells do not overlap. (E) The domain stretches over time, representing tissue growth, diluting chemoattractant and contributing to cell movement. A sample simulation at three time-points can be seen in (F-H). The grey contour represents where the chemoattractant gradient is too shallow to be sensed. (I) Effect of changing the number of leader cells on model migration - We varied the time after which newly inserted cells are followers instead of leaders. $n = 20$ repeated stochastic simulations for five different average leader fractions $\langle f_L \rangle$. On top of the overlays are greyscale boxplots of the distance migrated in the x-direction. The color of each boxplot shows the relative stream density (average number of cells divided by maximum over all parameter values). (J) A histogram of cell counts vs. distance migrated for five different leader fractions.

Figure 5. Trailblazer NC cells at the invasive front of the stream have a partially conserved molecular profile across distinct phases of migration. (A) Heat map of single trailblazer NC cells at HH Stage 13 ($n=72$ profiled from 356 cells isolated from 35 embryos) and Stage 15 ($n=76$ profiled from 212 cells isolated from 38 embryos) with

genes ordered by mean log2 expression in all cells based upon the 96-gene profile. (B) Principal component analysis (PCA) projections of 96 genes analyzed in single HH Stage 13 and 15 trailblazer NC cell profiling. Blue, shaded squares within each plot represent an arbitrary PC score (PC1-2) of less than 0.15. (B') and (B'') are magnified insets from PCA projections. (C) Violin plots of selected genes with a violin plot reference key.

Figure 6. Single trailblazer NC cells have a molecular profile that is distinct from the rest of the cranial NC stream. (A) Isolation of single NC cells (blue circle) from each quartile of the cranial NC stream. (B) PCA of single trailblazer and quartile NC cells. (C) Violin plots of selected genes. (D) Hierarchical clusterings of single trailblazer and quartile NC cells by either Euclidean distance or Pearson dissimilarity based upon averages of the 96-gene profiles. n=318 cells total; n=72 HH Stage 13 trailblazers, n=76 HH Stage 15 trailblazers, n=43 1st quartile, n=41 2nd quartile, n=44 3rd quartile and n=42 4th quartile.

Figure 7. Knockdown of key transcription factors perturbed cranial NC cell migration. (A) Cranial NC stream electroporated with fluorescein-labeled control morpholino (n=8 embryos), HAND2 morpholino (n=13 embryos) or TFAP2A morpholino (n=11 embryos) (green) and nuclear H2B mCherry (red) at HH St15. Asterisk is end of migrating cells into the brachial arches, arrow is thinnest portion of steam. (B-C) Distribution of the percentage of transfected NC cells along the migratory route at HH St13 and St15. (D-E) Box plot of the distance migrated at HH St13 and HH St15. (F-G) Box plot of the area covered at HH St13 and HH St15. Ba, branchial arch.

Figure 8. Upregulating leader behavior in trailing cell population perturbs stream migration and disrupts model migration. (A) Trailing NC cells transfected with control EGFP construct (n=20 embryos), HAND2 FL construct (n=13 embryos) or TFAP2A FL construct (n=14 embryos) (green) and Dil (red) at HH Stage 15. Asterisk is end of migrating cells into the brachial arches, arrow is thinnest portion of stream. (B) Distribution of the percentage of transfected NC cells along the migratory route. (C) Distribution of the average number of transfected NC cells along the migratory route. (D-E) To represent an upregulation of a leader-like behavior in the trailing portion of the stream, after time t_{LF} every other cell inserted into the domain is a leader cell (orange) rather than a follower cell (purple). (D) Overlay of 20 repeats of the stochastic simulation for control (WT) and perturbed (lead+) simulations. Greyscale boxplots show distance migrated in the x-direction and their color gives relative stream density (as in Fig. 4). (E) Histogram of cell counts vs. distance for control and perturbed simulations. r, rhombomere; ba, branchial arch.

Figure 9. Profile of trailblazer cranial neural crest cells. Molecular signature of trailblazer cranial neural crest cells. The molecular signature of trailblazers is comprised of 16 genes. These genes are highly and stably expressed in trailblazers overtime when compared to neural crest cells throughout the stream. All of the genes are regulated by the transcription factors TFAP2A and HAND2; arrows indicate shortest pathway regulation (direct, indirect, positive or negative) from Pathway Studios.

Table 1. Computational model parameters.

Table 2. Highest expressed genes in single trailblazer cells. Bolded genes are highly expressed in at least 50% of neural crest cells analyzed.

Table 3. Differential expression between trailblazers and quartiles at HH 15.

Bolded genes are highly expressed by trailblazers and differentially expressed between trailblazers and the rest of the stream.

Table 4. Molecular profile after HAND2 or TFAP2A knockdown. Bolded genes are either in the trailblazer molecular signature or HAND2 and TFAP2A.

MOVIE LEGENDS

Movie S1: Cell migration with low (~5%) vs high (> 50%) leader fraction. Newly inserted cells are followers from tLF = 6h (start of migration) and tLF = 11:25h, respectively.

Movie S2: Migration without leader cells. Without a directional cue, cells (grey) move in a random direction at each time step. Chemoattractant concentration does not affect cells and is shown for comparison with other simulations only.

SUPPLEMENTARY MATERIAL

Supplementary Figure 1. HCR analysis of mRNA expression in cranial neural crest cells at higher spatial resolution and at different phases of migration. (A)

40x images of DAPI staining, BAMBI HCR signal, HAND HCR signal, HNK-1 immunohistochemistry and overlay of the cells found in a representative branchial arch 2 (Quartile 1). (B) Box plots of the BAMBI and HAND2 HCR sum intensities from regions of interest (ROIs) from the different quartiles within the migratory stream at HH St. 15. Quartile 1 (blue), 2 (brown), 3 (purple) and 4 (pink). (C) 40x images of HNK-1 immunohistochemistry or FoxD3 HCR signal, BAMBI HCR signal, HAND2 HCR signal and overlay at HH Stages 9, 11 and 13. BAMBI expression is higher in lead neural crest at different stages of migration (white box) whereas HAND2 expression is expressed in very few cells.

Supplementary Figure 2. FACS is the most efficient single cell isolation method, and it does not significantly alter the gene expression profile. (A) Pairwise correlation of average log2 expression values produced by three distinct isolation methods. (B) Efficiency of successful single cell profiling as a percentage of the total number of single cells harvested for each method based upon outlier detection.

Supplementary Figure 3. Pairwise comparisons and single cell correlation intensity plot of trailblazer and quartile single cell sample groups. (A) Same description as in Supplementary Figure 2A. (B) Similarity matrix of each cell's 96-gene profile to all other single cell profiles as measured by Pearson's correlation coefficient. n=318 cells total.

Supplementary Figure 4: Violin plots of trailblazer and quartile single cell sample groups. (A) Violin plots of log2 gene expression in single NC cells isolated from HH13 Trailblazers (red), HH15 Trailblazers (green), the 1st quartile of the HH15 NC stream

(blue), the 2nd quartile of the HH15 NC stream (brown), the 3rd quartile of the HH15 NC stream (purple), and the 4th quartile of the HH15 NC stream (pink).

Supplementary Figure 5. Cell death analysis after morpholino transfection. (A) 10x images of neural crest cells transfected with control MO or HAND2 MO (purple) and stained with cleaved caspase-3 (green). (B) Average percentage of cells co-labeled with cleaved caspase-3 and morpholinos. White circles show examples of cleaved caspase-3 positive cells. n=6 embryos, 488 cells for control MO, n=6 embryos, 442 cells for HAND2. Error bars represent the SEM, p=0.5.

Supplementary Figure 6: Measurements of neural crest cell size and filopodial length derived for computational model parameters. Histograms of cell size and filopodia length measurements. Solid lines show mean, dashed lines show the median.

REFERENCES

- Adams, M. S., Gammill, L. S. and Bronner-Fraser, M.** (2008). Discovery of transcription factors and other candidate regulators of neural crest development. *Dev Dyn* **237**, 1021-33.
- Barrallo-Gimeno, A., Holzschuh, J., Driever, W. and Knapik, E. W.** (2004). Neural crest survival and differentiation in zebrafish depends on mont blanc/tfap2a gene function. *Development* **131**, 1463-77.
- Berg, H. C. and Purcell, E. M.** (1977). Physics of chemoreception. *Biophys J* **20**, 193-219.
- Brewer, S., Jiang, X., Donaldson, S., Williams, T. and Sucov, H. M.** (2002). Requirement for AP-2alpha in cardiac outflow tract morphogenesis. *Mech Dev* **110**, 139-49.
- Chazaud, C., Oulad-Abdelghani, M., Bouillet, P., Decimo, D., Chambon, P. and Dolle, P.** (1996). AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. *Mech Dev* **54**, 83-94.
- Choi, H. M., Beck, V. A. and Pierce, N. A.** (2014). Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* **8**, 4284-94.

Choi, H. M., Chang, J. Y., Trinh le, A., Padilla, J. E., Fraser, S. E. and Pierce, N. A. (2010). Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol* **28**, 1208-12.

D'Autreaux, F., Morikawa, Y., Cserjesi, P. and Gershon, M. D. (2007). Hand2 is necessary for terminal differentiation of enteric neurons from crest-derived precursors but not for their migration into the gut or for formation of glia. *Development* **134**, 2237-49.

Gallardo, V. E., Liang, J., Behra, M., Elkahoul, A., Villablanca, E. J., Russo, V., Allende, M. L. and Burgess, S. M. (2010). Molecular dissection of the migrating posterior lateral line primordium during early development in zebrafish. *BMC Dev Biol* **10**, 120.

Gammill, L. S. and Bronner-Fraser, M. (2002). Genomic analysis of neural crest induction. *Development* **129**, 5731-41.

Gammill, L. S. and Bronner-Fraser, M. (2003). Neural crest specification: migrating into genomics. *Nat Rev Neurosci* **4**, 795-805.

Hamburger, V. and Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. *Journal of Morphology* **88**, 49-92.

Hendershot, T. J., Liu, H., Sarkar, A. A., Giovannucci, D. R., Clouthier, D. E., Abe, M. and Howard, M. J. (2007). Expression of Hand2 is sufficient for neurogenesis and cell type-specific gene expression in the enteric nervous system. *Dev Dyn* **236**, 93-105.

Howard, M., Foster, D. N. and Cserjesi, P. (1999). Expression of HAND gene products may be sufficient for the differentiation of avian neural crest-derived cells into catecholaminergic neurons in culture. *Dev Biol* **215**, 62-77.

Knight, R. D., Javidan, Y., Nelson, S., Zhang, T. and Schilling, T. (2004). Skeletal and pigment cell defects in the lockjaw mutant reveal multiple roles for zebrafish tfap2a in neural crest development. *Dev Dyn* **229**, 87-98.

Knight, R. D., Nair, S., Nelson, S. S., Afshar, A., Javidan, Y., Geisler, R., Rauch, G. J. and Schilling, T. F. (2003). lockjaw encodes a zebrafish tfap2a required for early neural crest development. *Development* **130**, 5755-68.

Kulesa, P. M. and Fraser, S. E. (1998). Neural crest cell dynamics revealed by time-lapse video microscopy of whole embryo chick explant cultures. *Developmental Biology* **204**, 327-344.

Kulesa, P. M. and Fraser, S. E. (2000). In ovo time-lapse analysis of chick hindbrain neural crest cell migration shows cell interactions during migration to the branchial arches. *Development* **127**, 1161-72.

Kulesa, P. M. and Gammill, L. S. (2010). Neural crest migration: patterns, phases and signals. *Dev Biol* **344**, 566-8.

Kulesa, P. M., Teddy, J. M., Stark, D. A., Smith, S. E. and McLennan, R. (2008). Neural crest invasion is a spatially-ordered progression into the head with higher cell proliferation at the migratory front as revealed by the photoactivatable protein, KikGR. *Dev Biol* **316**, 275-87.

Li, W. and Cornell, R. A. (2007). Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev Biol* **304**, 338-54.

Luo, T., Lee, Y. H., Saint-Jeannet, J. P. and Sargent, T. D. (2003). Induction of neural crest in *Xenopus* by transcription factor AP2alpha. *Proc Natl Acad Sci U S A* **100**, 532-7.

McKinney, M. C. and Kulesa, P. M. (2011). In vivo calcium dynamics during neural crest cell migration and patterning using GCaMP3. *Dev Biol* **358**, 309-17.

McLennan, R., Dyson, L., Prather, K. W., Morrison, J. A., Baker, R. E., Maini, P. K. and Kulesa, P. M. (2012). Multiscale mechanisms of cell migration during development: theory and experiment. *Development* **139**, 2935-44.

McLennan, R. and Kulesa, P. M. (2007). In vivo analysis reveals a critical role for neuropilin-1 in cranial neural crest cell migration in chick. *Dev Biol* **301**, 227-39.

McLennan, R., Teddy, J. M., Kasemeier-Kulesa, J. C., Romine, M. H. and Kulesa, P. M. (2010). Vascular endothelial growth factor (VEGF) regulates cranial neural crest migration in vivo. *Dev Biol* **339**, 114-25.

Miller, C. T., Yelon, D., Stainier, D. Y. and Kimmel, C. B. (2003). Two endothelin 1 effectors, *hand2* and *bapx1*, pattern ventral pharyngeal cartilage and the jaw joint. *Development* **130**, 1353-65.

Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. and Tjian, R. (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev* **5**, 105-19.

Molyneaux, K. A., Wang, Y., Schaible, K. and Wylie, C. (2004). Transcriptional profiling identifies genes differentially expressed during and after migration in murine primordial germ cells. *Gene Expr Patterns* **4**, 167-81.

Morrison, J. A., Bailey, C. M. and Kulesa, P. M. (2012). Gene profiling in the avian embryo using laser capture microdissection and RT-qPCR. *Cold Spring Harb Protoc* 2012.

Nishiyama, C., Uesaka, T., Manabe, T., Yonekura, Y., Nagasawa, T., Newgreen, D. F., Young, H. M. and Enomoto, H. (2012). Trans-mesenteric neural crest cells are the principal source of the colonic enteric nervous system. *Nat Neurosci* **15**, 1211-8.

Pfisterer, P., Ehlermann, J., Hegen, M. and Schorle, H. (2002). A subtractive gene expression screen suggests a role of transcription factor AP-2 alpha in control of proliferation and differentiation. *J Biol Chem* **277**, 6637-44.

Piotrowski, T. and Baker, C. V. (2014). The development of lateral line placodes: taking a broader view. *Dev Biol* **389**, 68-81.

Ridenour, D. A., McLennan, R., Teddy, J. M., Semerad, C. L., Haug, J. S. and Kulesa, P. M. (2014). The neural crest cell cycle is related to phases of migration in the head. *Development* **141**, 1095-103.

Schorle, H., Meier, P., Buchert, M., Jaenisch, R. and Mitchell, P. J. (1996). Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature* **381**, 235-8.

Shen, H., Wilke, T., Ashique, A. M., Narvey, M., Zerucha, T., Savino, E., Williams, T. and Richman, J. M. (1997). Chicken transcription factor AP-2: cloning, expression and its role in outgrowth of facial prominences and limb buds. *Dev Biol* **188**, 248-66.

Simoës-Costa, M. and Bronner, M. E. (2013). Insights into neural crest development and evolution from genomic analysis. *Genome Res* **23**, 1069-80.

Simoës-Costa, M., Tan-Cabugao, J., Antoshechkin, I., Sauka-Spengler, T. and Bronner, M. E. (2014). Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. *Genome Res* **24**, 281-90.

Srivastava, D., Thomas, T., Lin, Q., Kirby, M. L., Brown, D. and Olson, E. N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat Genet* **16**, 154-60.

Tarbashevich, K. and Raz, E. (2010). The nuts and bolts of germ-cell migration. *Curr Opin Cell Biol* **22**, 715-21.

Teddy, J. M. and Kulesa, P. M. (2004). In vivo evidence for short- and long-range cell communication in cranial neural crest cells. *Development* **131**, 6141-51.

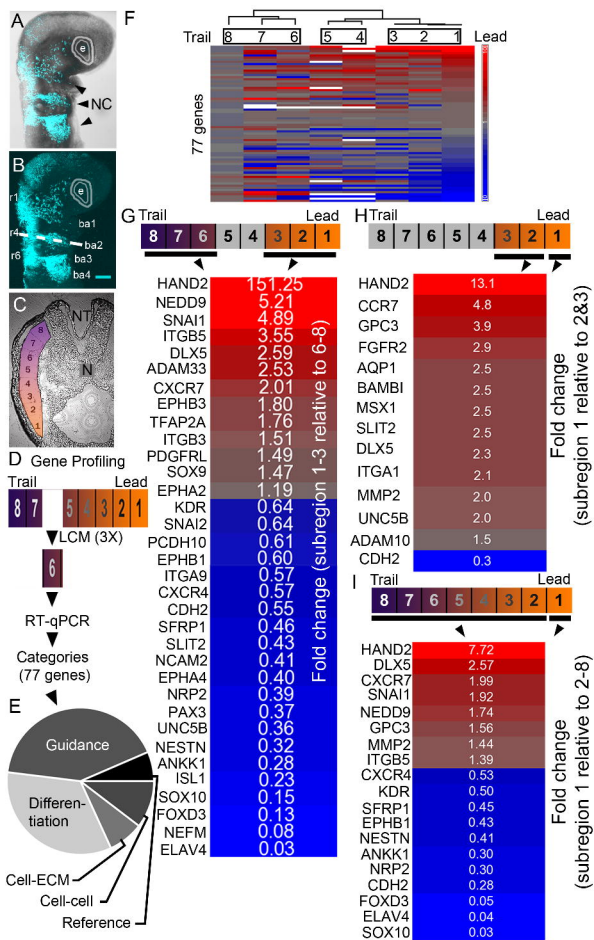
Thomas, T., Kurihara, H., Yamagishi, H., Kurihara, Y., Yazaki, Y., Olson, E. N. and Srivastava, D. (1998). A signaling cascade involving endothelin-1, dHAND and msx1 regulates development of neural-crest-derived branchial arch mesenchyme. *Development* **125**, 3005-14.

Voiculescu, O., Bodenstein, L., Lau, I. J. and Stern, C. D. (2014). Local cell interactions and self-amplifying individual cell ingression drive amniote gastrulation. *Elife* **3**, e01817.

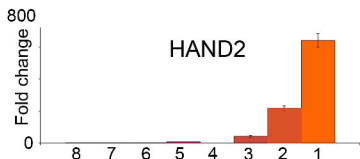
Wang, W. D., Melville, D. B., Montero-Balaguer, M., Hatzopoulos, A. K. and Knapik, E. W. (2011). Tfp2a and Foxd3 regulate early steps in the development of the neural crest progenitor population. *Dev Biol* **360**, 173-85.

Young, H. M., Bergner, A. J., Anderson, R. B., Enomoto, H., Milbrandt, J., Newgreen, D. F. and Whittington, P. M. (2004). Dynamics of neural crest-derived cell migration in the embryonic mouse gut. *Dev Biol* **270**, 455-73.

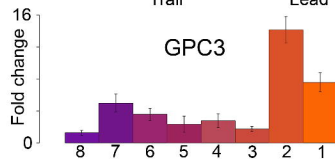
Young, H. M., Bergner, A. J., Simpson, M. J., McKeown, S. J., Hao, M. M., Anderson, C. R. and Enomoto, H. (2014). Colonizing while migrating: how do individual enteric neural crest cells behave? *BMC Biol* **12**, 23.



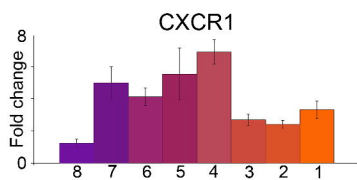
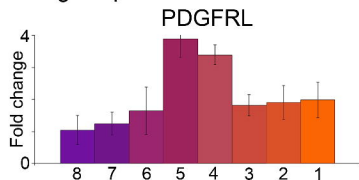
A High expression in the invasive front



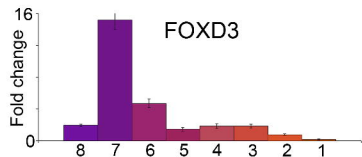
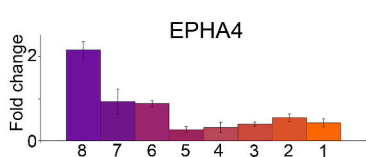
A' Key



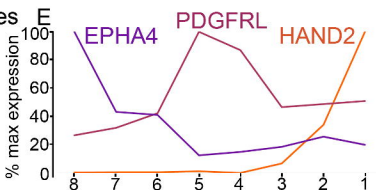
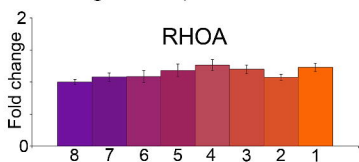
B High expression in middle of stream

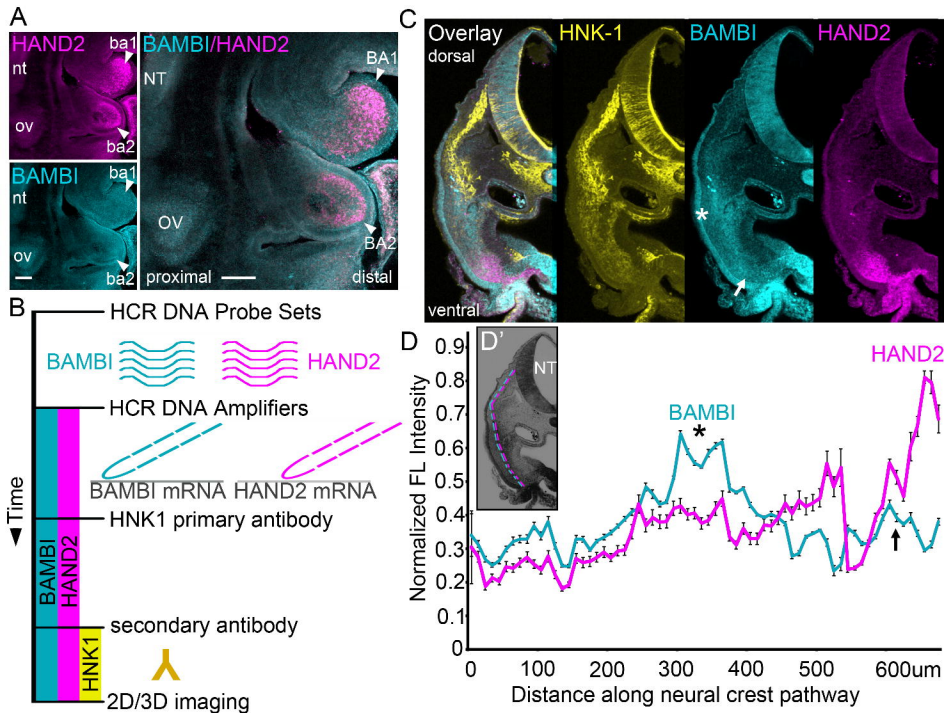


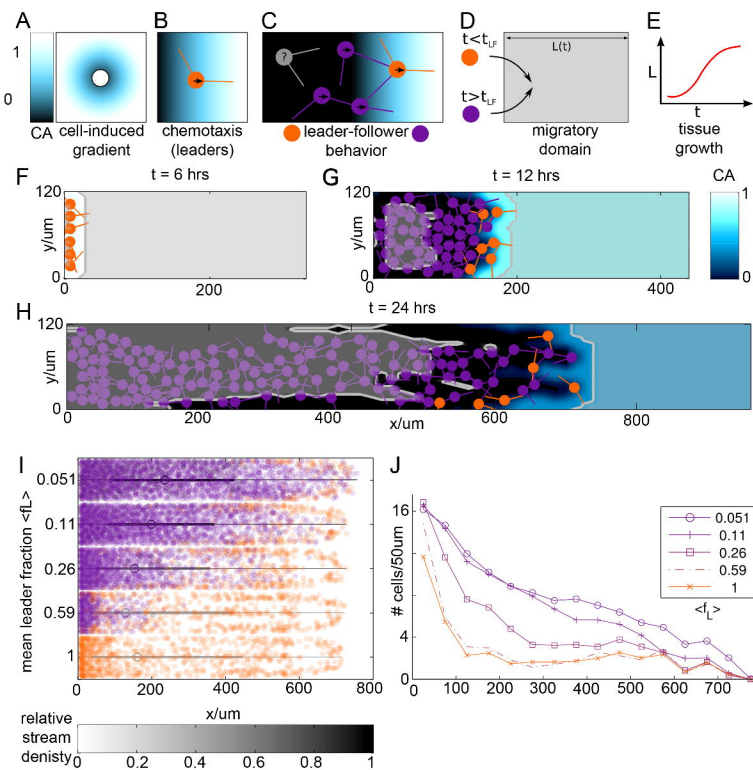
C High expression in trailing neural crest

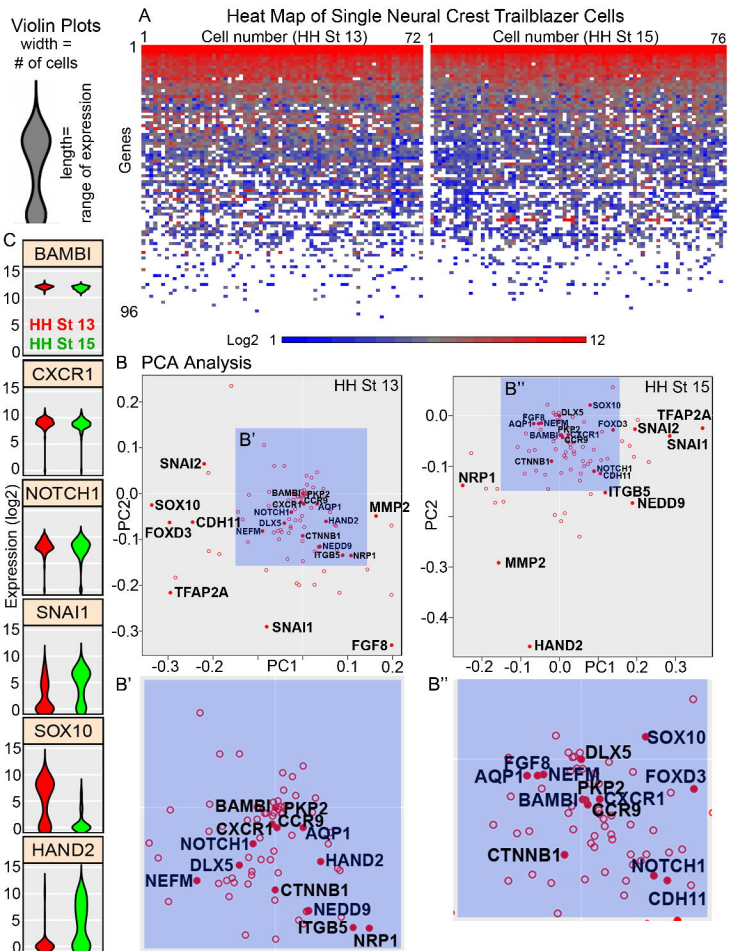


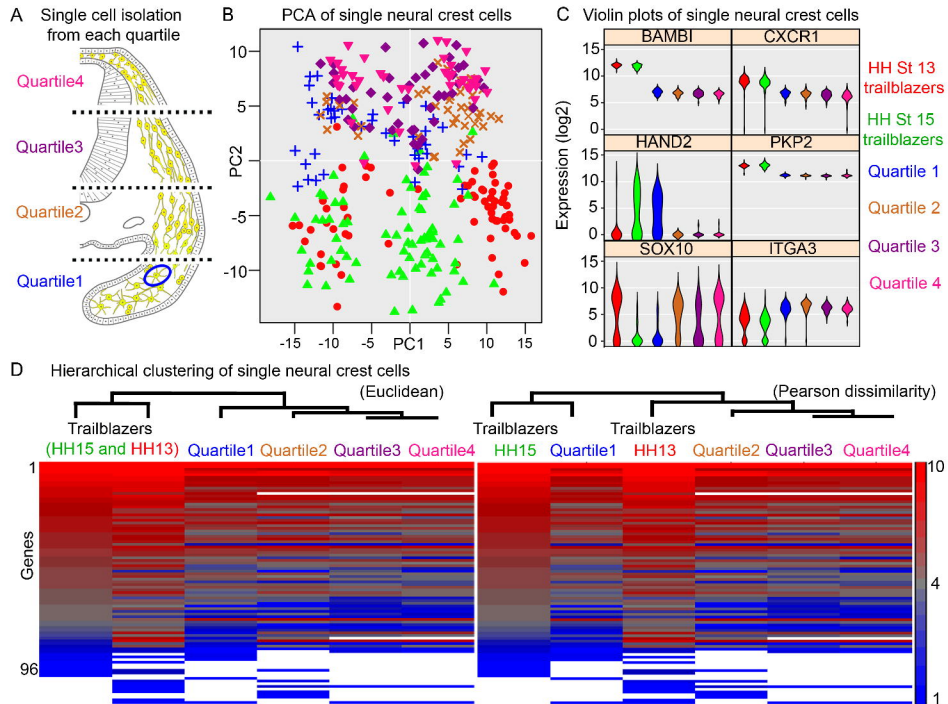
D No regional expression differences



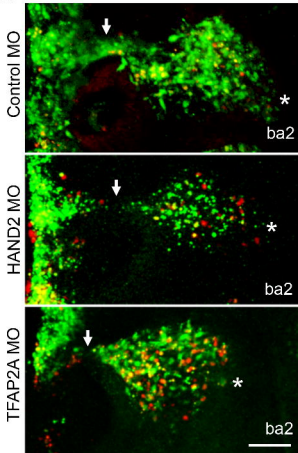




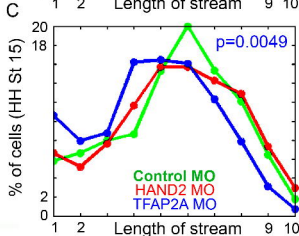
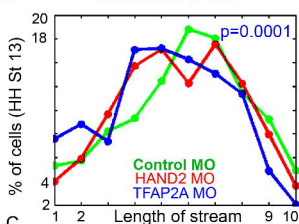




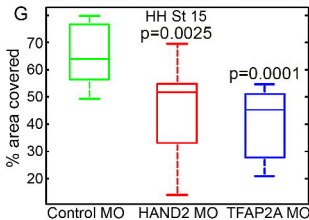
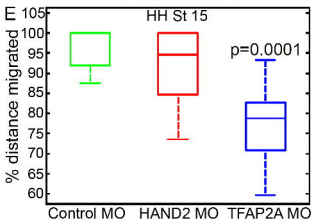
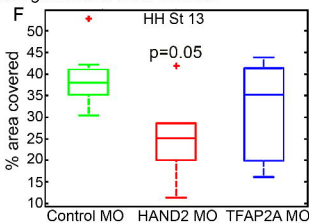
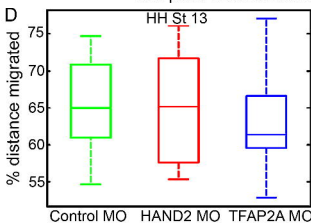
A Knockdown of HAND2 and TFAP2A



B Distribution of cells

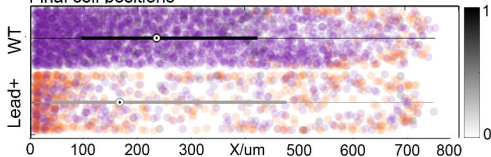


Box plots of cell distance migrated and area covered

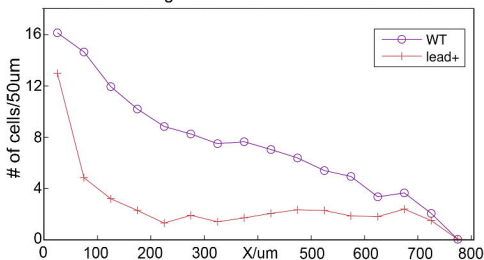


A Model simulations forcing lead behaviors in trailing cells

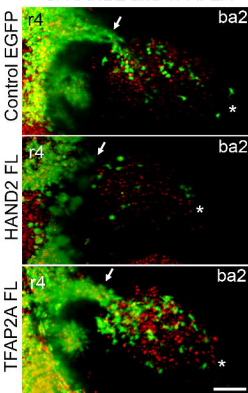
Final cell positions



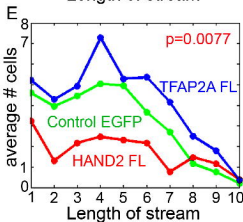
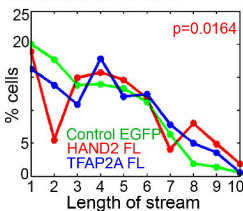
B Cell count histogram



C Overexpression of HAND2 and TFAP2A



D Distribution of cells



Trailblazer Molecular Signature

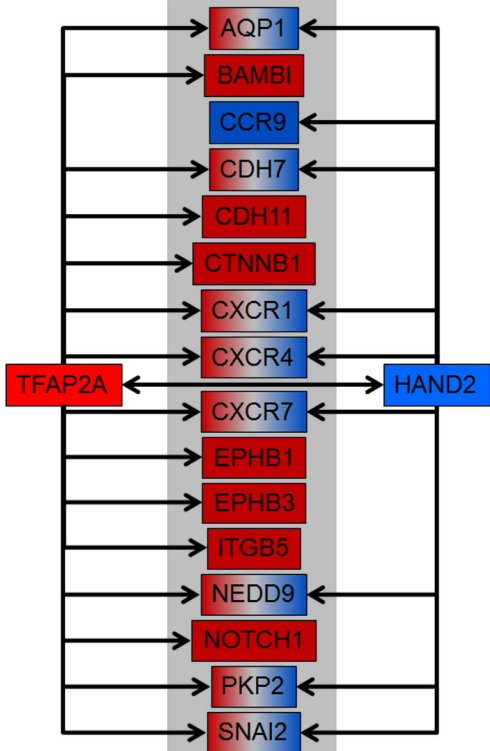


Table 1: **Model parameters**

Parameter values listed were used as a default, unless otherwise stated. Where a range is given, the model gives qualitatively similar results within that range, and the italicised value is the one used for the figures in this paper.

	Description	Value	Reference
t_{LF}	time after which newly inserted cells are followers	variable	see results section
n_{filo}	directions sampled per timestep	2	n/a, see notes
Δt	simulation time step	1 min	n/a
R	cell radius (nuclear)	$7.5\mu\text{m}$	McLennan & Kulesa (2010)
v_{lead}	cell speed (leader cells)	$41.6\mu\text{m/h}$	Kulesa <i>et al.</i> (2008)
v_{follow}	cell speed (follower cells)	$49.9\mu\text{m/h}$	Kulesa <i>et al.</i> (2008)
L_y	height of migratory domain	$120\mu\text{m}$	McLennan <i>et al.</i> (2012)
L_x	length of migratory domain (grows logistically, Eq. (1))	$300\mu\text{m}$ to $1100\mu\text{m}$	McLennan <i>et al.</i> (2012)
l_{filo}	sensing radius	$27.5\mu\text{m}$	see notes
$l_{\text{filo}}^{\text{max}}$	maximum cell separation before contact is lost	$45\mu\text{m}$	see notes
$\Delta c/c$	sensing accuracy	0.001 to <i>0.1</i>	Section S1.2
D	diffusion coefficient of chemoattractant (effective)	<i>0.1</i> to $10^5\mu\text{m}^2/\text{h}$	see notes
χ	production rate of chemoattractant	<i>0.0001</i> to $1/\text{h}$	see notes
λ	chemoattractant internalisation rate	100 to <i>1000</i> /h	see notes
k_{in}	rate at which cells enter the domain	10/h	see notes

Notes

Experimental time: Cell migration starts approximately six hours after electroporation ($t = 0$).

Directions sampled per timestep, n_{filo} : This cannot be directly related to the number of filopodia, which are greater in number, but sample at a lower speed (McLennan *et al.* 2012).

Diffusion coefficient of chemoattractant, D : The primary identified chemoattractant in chick cranial neural crest migration is VEGF¹⁶⁵ (McLennan *et al.* 2010). Its related isoform VEGF¹⁶⁴ is known to bind to extracellular matrix (ECM), and studies in angiogenesis estimate as little as 1% may be freely diffusing, the rest bound to ECM and cellular receptors (Mac Gabhann *et al.* 2006). Hence, we choose a low effective diffusivity.

Production rate of chemottractant, χ : In other tissues, VEGF production, or estimates thereof, range from 0.01-0.20 molecules/cell/s (Yen *et al.* 2011), $4.39\text{-}5.27 \cdot 10^{-5}$ molecules/ $\mu\text{m}^{-2}/\text{s}$ (Vempati *et al.* 2011) to $0.25 \cdot 10^{-17}$ pmol/ $\mu\text{m}^2/\text{s}$ (Mac Gabhann *et al.* 2006). In our system, the rate of VEGF production is unknown and difficult to measure. However, it is outweighed by internalisation through migrating neural crest cells, as VEGF is not seen to be replenished in trailing portions of the stream (McLennan *et al.* 2010). Thus, we assume χ to be low.

Chemoattractant internalisation rate, λ : To our knowledge, no estimates or measurements of VEGF internalisation rate of chick cranial neural crest exists. Angiogenesis studies have used values of $k_{\text{VEGFR2}} = O(10^{-4})/\text{s}$ (Mac Gabhann & Popel 2005, Yen *et al.* 2011). Berg & Purcell (1977) estimate the number of receptors needed for a near-optimal sensing accuracy as $N_R = R/s$, where R is the cell radius and s the receptor size. With $s = O(\text{nm})$, we can estimate the number of receptors to be $N_R \geq 10^4$. If receptor internalisation rates are comparable to other tissues, a lower bound on λ would be given by $k_{\text{VEGFR2}}N_R > 1$. However, the concentration of VEGF in our system is unknown, and hence the units of c , and therefore λ , in our model are arbitrary. We assume a high λ to ensure quick consumption of chemoattractant by cells.

Rate at which cells enter the domain, k_{in} : This is the rate of attempted cell insertions, in a typical simulation on the order of 10% of insertions are unsuccessful. It should be noted here again that our simulations are a two-dimensional abstraction of the three-dimensional migratory stream, which may contain 4-5 times as many cells *in vivo* in the transverse (z) direction.

Sensing radius, l_{filo} : This was calculated as the sum of the cell radius ($7.5\mu\text{m}$) and the mean filopodial length (which was directly measured from the cell body (Fig. S4B) to be $9\mu\text{m}$ and estimated from total cell size (Fig. S4A) to be circa $20\mu\text{m}$). Since we have only implemented contact between filopodium and cell body, but not between two filopodia, which does occur *in vivo* (Teddy & Kulesa 2004), we allow for a greater effective length.

Maximum cell separation before contact is lost, $l_{\text{filo}}^{\text{max}}$: The maximum cell size including filopodia was measured to be $86.3\mu\text{m}$ (Fig. S4A), half of which gives an estimate of maximum cell separation of $43.15\mu\text{m}$. Independent measurements of filopodial lengths gave a maximum of $30.4\mu\text{m}$ (from the cell body, Fig. S4B), which, together with the cell radius $R = 7.5\mu\text{m}$ and the average filopodial length (allowing for interflopodial contact) of $9\mu\text{m}$ gives an estimate of $46.5\mu\text{m}$.

Gene	AVE Ct of HH13 Trailblazers expressing	% of HH13 Trailblazers expressing
PKP2	13.00973678	100.00
BAMBI	13.99207538	100.00
CXCR1	16.923582	98.61
CXCR4	17.46313907	100.00
CTNNB1	17.51107514	100.00
CCR9	17.70692509	100.00
NOTCH1	17.73621817	98.61
CDH11	17.78358413	87.50
HAND2	17.88303114	6.94
FGF8	18.05853726	19.44
SOX10	18.42511568	68.06
FGFR1	18.48812662	94.44
SLIT1	18.50881211	88.89
TFAP2A	18.55281934	70.83
GPC3	19.12203206	100.00
ITGB3	19.15008624	73.61
ROBO1	19.20342733	100.00
ADAM10	19.27265348	98.61
EPHA4	19.31785986	86.11
FOXD3	19.37819436	62.50
SNAI2	19.48469388	75.00
CDH7	19.5414792	77.78
EPHB1	19.5593769	100.00
FZD7	19.59303938	98.61
ADAM33	19.77794749	79.17
AQP1	20.11626496	88.89
SFRP1	20.22819026	87.50
JAG1	20.34777223	69.44
CDH2	20.38084649	98.61
NRP2	20.41404755	90.28
KRT19	20.48268839	59.72
MMP2	20.56005797	55.56
NESTN	20.59934615	93.06
CXCL12	20.68492789	84.72
PAX3	20.700437	69.44
ITGB5	20.80221854	83.33
ERBB4	20.81292209	2.78
CXCR7	20.85370877	76.39
SNAI1	20.85531629	44.44
EPHA2	20.86312155	80.56
VEGFA	20.90223852	55.56
FGF4	20.99185856	8.33
TH	21.00074756	2.78
SPON1	21.00781815	76.39
CFC1B	21.10568255	95.83
PDGFRL	21.12772244	25.00
NRP1	21.17234955	38.89
EPHB3	21.17772359	93.06
NEDD9	21.18236349	54.17
SLIT2	21.22252072	52.78
BMPR2	21.29450543	87.50
PHOX2B	21.30238098	1.39
ITGA9	21.31944168	95.83
CCR7	21.33280342	47.22
PCDH10	21.35060574	12.50
ISL1	21.3965638	45.83
ITGA3	21.4559391	81.94
FGFR3	21.51790159	68.06
RUNX2	21.57578757	9.72
MMP9	21.57810155	9.72
ROBO2	21.62961883	79.17
ANGPT2	21.68445917	66.67
KDR	21.71726391	34.72
NEFM	21.76636245	63.89
EPHA3	21.82859112	20.83
CDH6	21.88794688	25.00
WISP1	21.91030979	1.39
EDNRA	21.92058669	76.39
ELAV4	21.94223617	2.78
VCAM1	21.97214993	1.39
NCAM2	22.25038961	6.94

Gene	AVE Ct of HH15 Trailblazers expressing	% of HH15 Trailblazers expressing
PKP2	12.93527783	100.00
BAMBI	14.22838368	100.00
CXCR1	17.20444133	98.68
CTNNB1	17.80330424	100.00
NOTCH1	17.88460284	96.05
CCR9	18.03276478	100.00
CXCR4	18.17937181	100.00
CDH11	18.43253169	96.05
FGF8	18.59746247	3.95
HAND2	18.8133826	59.21
FGFR1	18.84915588	98.68
TFAP2A	19.11212435	71.05
GPC3	19.29216546	100.00
ITGB5	19.62549784	93.42
ROBO1	19.66322555	100.00
CXCL12	19.80433937	94.74
SLIT1	19.80976563	85.53
MMP2	19.84402343	78.95
ADAM10	19.91648057	100.00
SNAI1	19.93185646	75.00
AQP1	20.04369513	82.89
SNAI2	20.18018349	81.58
JAG1	20.18576431	93.42
EPHB1	20.32304678	94.74
PCDH10	20.33257862	2.63
NEDD9	20.33689848	63.16
CDH7	20.46708801	67.11
ADAM33	20.47524681	80.26
ITGB3	20.49427049	88.16
FZD7	20.54024663	97.37
NRP1	20.62063313	57.89
NRP2	20.70717286	81.58
EPHA4	20.73164948	69.74
KRT19	20.79137561	47.37
CXCR7	20.80929894	59.21
EPHA2	20.82868704	85.53
KDR	20.93523081	43.42
CDH2	21.03293914	94.74
PDGFRL	21.12663188	53.95
VEGFA	21.18781963	48.68
ISL1	21.20050602	32.89
EDNRA	21.2195303	82.89
SPON1	21.318706	68.42
EPHB3	21.35982518	81.58
NESTN	21.3977451	84.21
CFC1B	21.41857232	93.42
NEFM	21.43616253	47.37
CCR7	21.44227401	65.79
SFRP1	21.49885016	71.05
BMPR2	21.50252218	81.58
SLIT2	21.55061038	47.37
UNC5B	21.56368985	80.26
ITGA9	21.56943449	89.47
VCAM1	21.66058977	9.21
MMP9	21.66456188	11.84
CDH6	21.77948169	21.05
SOX10	21.80649016	22.37
RUNX2	21.84936929	14.47
EPHA3	21.90526321	26.32
FGFR2	21.92818368	67.11
ANGPT2	21.94015081	68.42
FOXD3	21.98262749	40.79

Higher in trailblazers compared to quartile 1	p value	fold change
BAMBI	0.00E+00	16.82
PCDH10	9.52E-02	7.90
FGF8	1.40E-02	5.60
CCR9	0.00E+00	5.57
ELAV4	5.13E-02	4.89
FGFR2	5.93E-06	3.85
EPHB1	2.92E-06	3.66
NRP1	1.28E-03	3.34
CDH7	5.74E-07	3.01
MMP9	3.11E-02	2.94
PDGFRL	4.46E-06	2.84
AQP1	2.59E-04	2.66
CXCR1	4.35E-13	2.65
CDH6	1.16E-02	2.50
CXCR7	2.46E-04	2.32
SLIT1	1.83E-04	2.17
CXCR4	1.76E-09	2.15
PKP2	2.05E-12	2.13
CDH11	1.53E-05	2.08
CCR7	4.40E-03	2.08
CTNNB1	3.60E-08	1.81
ITGB5	5.98E-05	1.73
EPHB3	2.78E-03	1.72
SLIT2	3.36E-02	1.64
SNAI2	3.55E-02	1.60
NEDD9	3.37E-02	1.57
NOTCH1	3.47E-02	1.49

Higher in trailblazers compared to quartiles 2-4	p value	fold change
BAMBI	0.00E+00	20.60
HAND2	2.21E-05	18.53
PCDH10	3.64E-02	6.02
NRP1	5.28E-07	4.54
CCR9	0.00E+00	4.21
AQP1	7.77E-15	3.24
NEDD9	4.96E-10	3.21
CXCR1	0.00E+00	3.10
CXCL13	7.27E-02	3.02
ITGB5	3.04E-13	3.01
PDGFRL	5.96E-11	2.83
JAG1	6.45E-07	2.83
MMP2	2.67E-07	2.78
EPHA1	3.52E-04	2.76
FGFR2	1.13E-06	2.64
CXCR7	1.26E-06	2.52
SLIT2	2.30E-06	2.42
PKP2	0.00E+00	2.22
ISL1	2.49E-02	2.11
RUNX2	4.30E-02	2.06
NEFM	2.59E-03	2.01
KDR	9.39E-02	1.99
NOTCH1	7.14E-07	1.99
CTNNB1	0.00E+00	1.97
CXCR4	4.29E-11	1.87
EPHB3	2.10E-05	1.83
SOX9	3.86E-02	1.81
EPHB1	1.15E-04	1.81
CCR7	2.83E-03	1.68
CDH11	6.32E-05	1.62
CXCL12	1.57E-03	1.57
CDH6	9.43E-02	1.48
CDH7	1.49E-02	1.41
UNC5B	4.22E-02	1.40
BMP2	6.78E-02	1.28
SNAI2	3.68E-02	1.28
ROBO1	6.06E-02	1.16

Gene	p	Higher in HAND2 MO	Lower in HAND2 MO
HAND2	0.06899	2.17	
EPHA6	0.08247	2.023	
CDH7	0.07705	0.723	1.383125864
CDH6	0.04665	0.692	1.445086705
CDH11	0.03969	0.529	1.890359168
ITGA9	0.05754	0.696	1.436781609
ITGB3	0.01035	0.542	1.84501845
TFAP2A	0.09821	0.657	1.522070015
SPON1	0.03555	0.574	1.742160279
SOX10	0.02956	0.479	2.087682672
PAX3	0.02137	0.416	2.403846154
FGFR2	0.08873	0.846	1.182033097
FGFR1	0.03253	0.734	1.36239782
EPHB1	0.01072	0.676	1.479289941
NRP2	0.02771	0.593	1.686340641
CXCR4	0.04437	0.292	3.424657534
AQP1	0.07819	0.301	3.322259136
WISP1	0.05466	0.23	4.347826087
KDR	0.0691	0.347	2.88184438

Gene	p	Higher in TFAP2A MO	Lower in TFAP2A MO
HAND2	0.002251	3.764	
EPHA6	0.08972	2.683	
MMP2	0.01251	2.165	
ITGB5	0.02536	1.471	
NEDD9	0.08225	1.321	
FGFR2	0.08381	1.067	
EPHB1	0.05839	0.703	1.422475107
CXCR4	0.06061	0.4	2.5
WISP1	0.07412	0.345	2.898550725