Data Article

Transcriptome dataset of trunk neural crest cells migrating along the ventral pathway of chick embryos

Christina Murko\textsuperscript{a,b,*}, Felipe Monteleone Vieceli\textsuperscript{a}, Marianne Bronner\textsuperscript{a}

\textsuperscript{a} Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
\textsuperscript{b} Center for Anatomy and Cell Biology, Medical University of Vienna, Vienna, Austria

\textbf{Abstract}

We present a transcriptome dataset generated from migratory chick trunk neural crest cells, which are destined to form components of the peripheral nervous system. Using the Sox10E1 enhancer, which specifically labels neural crest cells migrating on the trunk ventral pathway, we performed fluorescence activated cell sorting (FACS) of electroporated embryos to obtain a pure population of these cells for library preparation and Illumina sequencing. The results provide a list of genes that are enriched in the trunk neural crest. To validate the data, we performed in situ hybridization to visualize expression of selected transcripts.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\section*{Specifications table}

\begin{tabular}{|l|l|}
\hline
Subject area & Biology \\
More specific subject area & Developmental Biology \\
Type of data & Figures, Graphs, Tables \\
\hline
\end{tabular}

* Corresponding author.

E-mail addresses: christina.murko@meduniwien.ac.at (C. Murko), fvieceli@caltech.edu (F.M. Vieceli), mbronner@caltech.edu (M. Bronner).

https://doi.org/10.1016/j.dib.2018.11.109

2352-2409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
How data was acquired | Data was obtained with Illumina HiSeq. 2000. RNA in situ hybridization was used to validate gene expression
---|---
Data format | Analyzed
Experimental factors | Embryos were electroporated to label migrating neural crest cells, which were then isolated by fluorescent activated cell sorting (FACS).
Experimental features | Electroporated, sorted embryonic cells
Data source location | Pasadena, USA
Data accessibility | Data are within the article and provided as Supplementary files. Raw sequencing data have been deposited into the NCBI sequence read archive (SRA) under BIO Project PRJNA494045
Related research article | Simões-Costa et al. [1]

**Value of the data**

- The data will be useful for other researchers in the field to determine specific gene expression levels in the ventrally migrating trunk neural crest population.
- The data allow comparison with previously published datasets and provides gene expression patterns within subpopulations of neural crest cells.
- This screen could contain new targets for the treatment of diseases arising from the trunk neural crest (e.g. neuroblastoma, Hirschsprung disease).

### 1. Data

The neural crest is a multipotent embryonic cell population that gives rise to many different cell types in vertebrates. Initially specified in the dorsal neural tube, these cells soon detach by undergoing an epithelial to mesenchymal transition and start to migrate throughout the embryonic body. The developmental potential of neural crest cells differs according to their axial level of origin and the migratory route they follow. Hence, neural crest cells on each pathway exhibit distinct gene expression profiles [2].

Our group has utilized enhancer based cell sorting together with next generation sequencing to generate genome-wide expression profiles of specific neural crest subpopulations in the chicken embryo [1,3,4]. Here, we present the dataset we have generated from neural crest cells migrating along the trunk ventral pathway. To label these cells, we have used a GFP reporter driven by the Sox10E1 enhancer that is specifically active in neural crest cells following this pathway [5,6]. The cells were isolated by FACS and sequenced on an Illumina HiSeq platform. By comparing the sorted GFP+ cell population with unlabeled cells from the rest of the trunk region (GFP-), we have obtained a list of transcripts that are enriched in the neural crest.

The data show enrichment of genes associated with neurogenesis and cell migration as well as a number of signaling pathways including Wnts and BMPs (bone morphogenetic protein) (Fig. 1 and Supplementary lists). We have further validated the data by performing whole mount in situ hybridization of selected transcripts (Fig. 2, Fig. S1 and Table 1). For that purpose, we selected transcripts that already had known neural crest expression (e.g. TFAP2A), transcripts that have not been analyzed in the neural crest but have been shown to play important roles in processes that may be relevant to neural crest migration or differentiation or are part of enriched signaling pathways (e.g. DTX4), and transcripts that have not been described in any related context so far (e.g. AGPAT4).
2. Experimental design, materials, and methods

2.1. Embryos

Fertilized chicken eggs were obtained from McIntyre Poultry & Eggs (Lakeside, CA) and incubated at 37 °C until embryos reached the desired developmental stage. Embryos were harvested in Ringer’s solution and staged according to the criteria of Hamburger and Hamilton [7]. Embryos for in situ hybridization were fixed in 4% paraformaldehyde (PFA) overnight at 4 °C, washed in phosphate buffered saline (PBS), dehydrated in a methanol series and stored at -20 °C until further processing.
2.2. Electroporation

Embryos were electroporated in ovo as previously described [8]. A pTKGFP vector containing the Sox10E1 enhancer sequence (described in 6) was injected at 2 μg/μl into the neural tube lumen of HH13–14 (2 days old) embryos, followed by electroporation with platinum wire electrodes for 3 pulses of 20 V for 30 ms, with an interval of 100 ms between pulses. This was followed by another round of electroporation with reversed polarity to achieve DNA uptake on both sides of the neural tube. Eggs were then sealed and reincubated until the embryos reached stage HH18–19 (3 days).

2.3. Cell dissociation and FACS

Electroporated embryos that exhibited robust GFP expression were harvested and washed in sterile PBS. Using a fluorescent stereoscope, the trunk was cut from the level of the wing bud until the leg buds. Tissue samples were pooled together and cells were dissociated with Accumax (Accutase SCR006). Clumps of cells were removed by passing through a 40 μm cell strainer (BD Biosciences). GFP+ and GFP- cells were sorted using a BD FACS ARIA Cell Sorter (BD Bioscience). The 7- AAD viability dye (Thermo Fisher) was used to exclude dead and damaged cells. We manually choose the cut off for sorting the GFP+ cell population at stringent conditions to minimize contaminations with false positive auto-fluorescent cells (e.g. blood cells).

2.4. Library preparation and sequencing

We used the RNAqueous Micro kit (Ambion) to isolate RNA from ~50,000 sorted cells per replicate from each GFP+ and GFP- populations. RNA quality was assayed in an Agilent 2100 Bioanalyzer, and only
samples with an RNA integrity number (RIN) > 8 were further processed. Concentration of the samples was measured with a Qubit fluorometer. 50 ng of RNA were used for RNA amplification and cDNA synthesis with the Ovation RNA-Seq System V2 (NuGEN). 2 μg of cDNA were then used for generating SR50 datasets with at least 40 million reads depth per sample in a HiSeq. 2000 Illumina instrument. 2 biological replicates were used for each condition (GFP+ and GFP-).

2.5. RNA-seq analysis

Sequenced reads were aligned to the *Gallus gallus* genome (galGal5.91) with HISAT2 [9], counted with featureCounts [10] and compared with DESeq. 2 [11] to identify differentially expressed transcripts. Statistical significance of genes that are enriched in the neural crest population is based on a 0.05 false discovery rate (Benjamini-Hochberg method). Functional classification of genes enriched in the neural crest was performed with the ToppGene suite [12] and data from the Gene Ontology Consortium [13], PantherDB [14], and DisGeNET [15].

| Gene name | Gene ID | P-adj  | log2(FC) | in situ |
|-----------|---------|--------|----------|--------|
| DLL1      | ENSGALG00000011182 | 7.04122E-22 | 2,546658444 | NC     |
| SOX2      | ENSGALG000000043460 | 8.32583E-15 | 2,275654915 | NC     |
| AGPAT4    | ENSGALG00000011571 | 0.007149007 | 1,345032181 | NC     |
| ANKFN1    | ENSGALG00000003105 | 5,59869E-08 | 1,959835002 | NC     |
| HES5      | ENSGALG00000001141 | 0.007426558 | 1,556181559 | NC     |
| TALLN3    | ENSGALG000000015379 | 0.012212793 | 1,62649238 | NC     |
| DTX4      | ENSGALG00000010835 | 1,43041E-09 | 1,833313017 | NC     |
| PAX7      | ENSGALG00000042204 | 8,07397E-16 | 2,169404336 | NC     |
| HES6      | ENSGALG00000028415 | 2,331E-16  | 2,309646573 | NC     |
| DRAxin    | ENSGALG00000004631 | 8,07397E-16 | 2,409119463 | NC     |
| PAX3      | ENSGALG000000030944 | 3,09434E-20 | 2,046728588 | NC     |
| SOX10     | ENSGALG00000012290 | 2,24326E-14 | 1,64523744 | NC     |
| MOXD1     | ENSGALG00000002911 | 0.000159845 | 1,01172748 | NC     |
| TFAP2A    | ENSGALG00000012775 | 0.041892446 | 0,860921861 | NC     |
| SCRT2     | ENSGALG00000028912 | 2,78475E-08 | 2,52286606 | NC     |
| RHBDL3    | ENSGALG00000003426 | 4,77513E-13 | 2,36586812 | NC     |
| GDF10     | ENSGALG00000005985 | 1,93411E-07 | 2,296980425 | NC     |
| H53T6     | ENSGALG00000005413 | 0,015486417 | 1,570936264 | NC     |
| PROX1     | ENSGALG00000009791 | 6,35622E-10 | 2,114807597 | NT     |
| NOTCH1    | ENSGALG00000002375 | 2,74837E-15 | 1,572570866 | NC     |
| DBX1      | ENSGALG00000003965 | 2,78759E-44 | 3,150177224 | NT     |
| ISM1      | ENSGALG000000009042 | 1,62958E-12 | 2,919884059 | NC     |
| SOX3      | ENSGALG00000040383 | 3,87094E-10 | 2,221184752 | NT     |
| FAM222A   | ENSGALG00000029544 | 0,000177813 | 1,799084864 | NT     |
| RFX4      | ENSGALG00000012647 | 1,26809E-14 | 2,027190365 | NT     |
| SOX13     | ENSGALG00000000583 | 7,59841E-09 | 1,93016488 | NT     |
| PAX6      | ENSGALG00000012123 | 4,70928E-23 | 2,502697167 | NT     |
| WNT4      | ENSGALG000000041708 | 3,74752E-30 | 2,544362001 | NT     |
| CHL1      | ENSGALG000000037856 | 4,29E-21  | 2,35611689 | NT     |
| FZD3      | ENSGALG00000042308 | 2,12E-13  | 1,847532482 | NT     |
| LRRC4C    | ENSGALG00000007948 | 2,28E-07  | 1,693090154 | NT     |
| PTPRN2    | ENSGALG00000030054 | 2,10E-10  | 1,817486577 | NT     |

Summary of all transcripts that were verified by in situ hybridization. Abbreviations: NT: dorsal neural tube (premigratory neural crest); NC: migrating neural crest.
2.6. Probe preparation and in situ hybridization

Dioxigenin labeled antisense probes were generated by in vitro transcription using Promega RNA Polymerases. Probes were either generated from linearized template DNAs or directly amplified by PCR and addition of T7 recognition sites to the antisense primers. The following probes were used:

- **DLL1**: described in [16];
- **HES6**: linearized template made from ChEST clone 62d7;
- **SOX2**: described in [17];
- **TFAP2A**: linearized template made from ChEST clone 401g22;
- **MOXD1**: described in [18].

The following probes were amplified by PCR using gene specific primers:

- **DTX4**: Fwd: CATCGGCTTCTGCTACGTGA, Rev: taatacgactcactataggAGACCAGTCGGATGTACCA;
- **AGPAT4**: Fwd: TGGACATCGTTGGCTTTCTGA, Rev: taatacgactcactataggATAGGCACTGCTGGGTAGGT;
- **TAGLN3**: Fwd: GGCAAGCATTAGAGATGGCT, Rev: taatacgactcactataggCGCCTCAGAGCACTAACTAT

Whole mount RNA in situ hybridization was performed as described previously [19]. Post hybridization washes were carried out using MABT (maleic acid buffer containing 0.1% Tween-20) and target-specific probe binding was visualized using NBT/BCIP (nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolylphosphate p-toluidine salt) as the color substrate. The time to develop the color reaction was probe specific and varied from 2 h to several days. A mixture of embryos at developmental stages between HH16 and HH20 was used for each probe to assess the spatio-temporal gene expression pattern of each transcript during the course of migration (e.g. from premigratory/delaminating to migratory and coalescence of dorsal root ganglia).

Acknowledgments

We thank Diana Perez and Igor Antoshechkin for help with FACS and sequencing. This work was supported by a postdoctoral fellowship from the Curci foundation and an Erwin Schrödinger fellowship from the Austrian Science Fund FWF (J3538-B19) to CM and NIH grant R01DE024157 to MEB.

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.11.109.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.11.109.

References

[1] M. Simões-Costa, J. Tan-Cabugao, I. Antoshechkin, T. Sauka-Spengler, M.E. Bronner, Transcriptome Analysis reveals novel players in the cranial neural crest gene regulatory network, Genome Res. 24 (2014) 281–290. https://doi.org/10.1101/gr.161182.113.

[2] M. Rothstein, D. Bhattacharya, M. Simões-Costa, The molecular basis of neural crest axial identity, Dev. Biol. (2018), https://doi.org/10.1016/j.ydbio.2018.07.026 (30598-5).
