UC Riverside

UCR Honors Capstones 2020-2021

Title

A COMPARATIVE TRANSCRIPTOMIC ANALYSIS OF DIFFERENTIATING NEURAL CREST CELLS AND NEUROECTODERM OR ECTODERM CELLS AT EARLY DEVELOPMENT

Permalink

https://escholarship.org/uc/item/1cp610mk

Author

Armstrong Suthahar, Sujit Silas

Publication Date

2021

Data Availability

The data associated with this publication are within the manuscript.

A COMPARATIVE TRANSCRIPTOMIC ANALYSIS OF DIFFERENTIATING NEURAL

CREST CELLS AND NEUROECTODERM OR ECTODERM CELLS AT EARLY

DEVELOPMENT

By

Sujit Silas Armstrong Suthahar

A Capstone Project Submitted for Graduation with University Honors

March 04, 2021

University Honors
University of California, Riverside

APPROVED

Dr. Martín I. García-Castro
Department of Biomedical Sciences

Dr. Richard Cardullo, Howard H Hays Jr. Chair
University Honors

ABSTRACT

Neural Crest (NC) are thought to be derived from the ectoderm, yet they display a wider differentiation potential which includes cell types normally assigned to a mesodermal origin. Our study compares the transcription profiles of early-stage differentiation in neuroectoderm and neural crest from pluripotent embryonic stem cells. We aim to identify critical transcripts which are both common and different between these two fates, to assess if human NC generated from pluripotent stem cells truly emerges from neuroectodermal precursors, and when and how they acquire or retain transcripts associated with their broader differentiation potential. The revolutionizing advent of next generation sequencing technologies. have accelerated the acquirement of expression profiles of cell differentiation data. Our group has established a model based on human pluripotent stem cells (hESC), that generates NC cells in just 5 days that express markers and display expected NC-differentiation potential. We further generated transcriptome data of the transition from hESC to NC cells, and here we aim to compare these profiles of expression to those generated from other expression profiles that characterize the generation of neuroectoderm from human pluripotent stem cells. Literature reports of outstanding genes identified in our analysis will be used to support the requirement or contribution of these genes to specific fates (NC and NE). Limma voom gene expression analysis tool along with other relevant bioinformatics tools will be used to obtain gene expression data. The overall goal is to identify a few genes necessary for NE that are expressed or not expressed in NC precursors.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

INTRODUCTION

Neural crest (NC) cells are a multipotent cell type unique to vertebrates, that emerges from the neural tube, migrate through stereotypic pathways, and contributes to a wide variety of cell types including neurons and glia of the peripheral nervous system, melanocytes of the skin, and craniofacial ectomesenchyme (bone and cartilage amongst others). They are known to be of ectoderm origin, and they are identified to be developed between the neural plate and the non-neural ectoderm. About one-third of congenital birth defects are caused due to improper NC differentiation (Prasad *et al.,* 2019). Numerous health conditions are caused by improper NC differentiation and include craniofacial defects (like cleft lip/palate), rare syndromes (Waardenburg syndrome), and deadly cancers (like melanoma). A better understanding of NC development is likely to aid in clinical efforts to ameliorate the effects of these taxing conditions.

The craniofacial contribution of NC to form bone, cartilage, adipose tissue, tooth forming cells, and other ectomesenchymal derivatives defy the canonical contributions of ectoderm, and thus the sequential restriction of potential (Prasad et al., 2019).

Neural Crest (NC) cells are believed to be derived from the ectoderm, which is responsible for the development of the skin or non-neural ectoderm, and neurons and glia of the central nervous system. In agreement, the NC generates peripheral neurons and glia, as well as melanocytes of the skin. However, cranial NC also contributes to the development of bone, cartilage, and adipose tissue which are normally thought to be derived from the mesoderm (Prasad et al., 2020). Their contribution to these mesodermal derivatives defies the current assumptions of sequential restriction of potential and doesn't comply with the classic germ layer theory. According to Waddington's Epigenetic landscape representing the process of cellular decision-making during development, cells can take specifically permitted trajectories, leading to

different outcomes or cell fates. As these cells take some paths, they sequentially limit their capacity to adopt other paths. This process broadly known as sequential segregation of potential is a current tenant during development (Ladewig et al., 2013; Waddington, 1942). However, NC cells seem to bypass this sequential segregation of potential, being an ectoderm derivative, they should have lost mesodermal capacities, yet, unlike other ectodermal derivatives, they display ectodermal and mesodermal potential.

According to recent research done on NC cells, it is evident that its origin can be traced to the early stages of embryogenesis preceding the establishment of the neural plate (Prasad et al., 2019). Research in chick embryos has identified that chromatin elements regulating NC related genes are accessible in the early epiblast stage and premigratory NC stage, defining a regulatory landscape that is established during gastrulation (Williams et al., 2019).

 In recent years, next-generation sequencing approaches have been widely used to study transcriptomes of various genomes (Qin, 2019). Considering the advantages of this approach, we aim to compare the gene expression data of transcriptomics obtained from the early stages of NE and NC differentiation. Our goal is to identify critical transcripts which are both common and different between these two fates, to assess if human NC generated from pluripotent stem cells truly emerges from neuroectodermal precursors, in which case a core of neuroectoderm genes critical for neuroectoderm development should be observed in NC development. Similarly, gene profiles associated with endo-mesodermal potential should be lost during these early facets. And later on, it would be expected for NC but not for NE to gain expression of key genes associated with mesodermal capacity endowing them with bone-cartilage potential. Alternatively, if NC is not derived from the ectoderm, one would expect to observe a retention of NE and mesodermal gene expression in NC precursors.

Literature reports of outstanding genes identified in our analysis will be used to support the requirement or contribution of these genes to specific fates (NC and neuro-ectoderm). We will use chromatin accessibility data from early embryogenesis to support this assumption further, with NC and NE differentiation to further identify a unique set of genes that are specific to each lineage. Connecting back to the idea of Waddington Epigenetics, NC and NE differentiation are presumably controlled by epigenetic regulations at early embryogenesis which gives NC cells a multipotent potential to form mesoectodermal derivatives.

In this analysis, we use transcriptomic data from Li et al. (NE differentiation data) and Prasad et al. (5-day protocol NC differentiation data) to identify gene overlaps between NC cells and NE cells. RNA-seq data of the early differentiation stage (day 2) of each of the studies were compared to find overlapping genes. Chromatin accessibility data from the research of William et.al was used to compare ATAC-seq open chromatin region data with early NC and NE differentiation. Other highly relevant data and manuscripts with publicly accessible experimental data were used to compare the early developmental stages. Gene Ontology enrichment analysis tool was used to find enriched genes. The differentially expressed genes were identified using the limma voom differential gene expression analysis tool on DEApp.

METHODOLOGY

Obtaining Relevant Experimental Data:

NE differentiation data was obtained from Li et al. reposited under the GEO accession ID

GSE103715. The standard ESC cultures were supplemented with fibroblast growth factor (FGF).

To achieve NE induction, the human ES cells were floated in an ES cell growth medium without

FGF. They were subsequently floated in the neural induction medium consisting of F-

12/DMEM, N2 supplement, and other non-essential amino. The NC differentiation data obtained

from Prasad et al. uses a 5-day protocol to induce NC cells (Leung et al., 2016). To achieve NC

induction, WNT/ β-catenin signaling was modulated in hPSCs. The time points selected for the

analysis were day 0 and day 2 for prospective neural crest differentiation and day 0 and day 2 for

the prospective NE differentiation. ATAC-seq chromatin accessibility data reposited under the

GEO accession ID GSE121318 was used in comparison to the early NC and NE differentiation.

Data for the open chromatin region reads of whole chick epiblast at Hamburger Hamilton stage

(HH4) were reposited under the accession ID.

Genome ID conversion:

Datasets with ENSEMBL gene IDs were converted to gene symbols for a better understanding of

the dataset. The gene ID converter tool from BioTools.fr was utilized to obtain gene symbols by

inputting the ENSEMBL IDs (*Dashboard - Genomics Biotools*, n.d.).

Limma Voom Differential Expression Analysis

DEApp, an open-source web application for differential expression analysis was used on RNA-

seq data from the studies. DEApp uses R program to run its codes and it effectively utilizes

Bioconductor packages such as DESeq, limma voom, and edgeR to run the differential

expression analysis. The program is hosted on a background cloud server which helps compute

next-generation sequencing data using any operating system. The program enables users to view the metadata and the counts data in a dynamic interface (Yan Li & Andrade, 2017).

The differential expression analysis involves a four-step process that included raw data and metadata input, data normalization, differential expression analysis, methods, and comparison. We used limma voom to perform the differential expression analysis on the RNA-seq counts data obtained from the studies. Limma voom uses a linear model to read RNA-seq counts data. It estimates the mean-variance relationship of log counts to generate a weight for each observation and enters them into a Bayes analysis pipeline (Law et al., 2014).

Once the RNA-seq counts data was run through limma voom, a volcano plot with the differentially expressed gene was obtained. The list with the top hits for upregulated and down-regulated genes were generated, and a multidimensional scaling plot was obtained to visualize the level of similarity or difference between the datasets.

The criteria set to identify the top upregulated and downregulated differentially expressed genes were p-values (< 0.05) and log fold change values (>1.5) for upregulated genes or (< -1.5) for downregulated genes.

Organization Gene Expression Data in an Excel File:

The gene lists with the expression data obtained from the transcriptomic studies and other relevant studies were transferred to Microsoft Excel spreadsheets and were arranged in the order of decreasing and increasing log fold change. The top upregulated (log fold change >1.5) and down-regulated (log fold change < -1.5) genes were isolated and were added to a new sheet to identify overlaps. Top 25 upregulated and downregulated genes unique to each data set with no overlaps were also identified. All the data corresponding analysis will be provided in a supplemental Excel file.

Identifying Overlaps Using Biotools.fr

Overlaps between genes were identified by using the Venn diagram tool at biotools.fr. The tools compared the gene symbols from the data to find overlaps (*Venn Diagrams - Genomics Biotools*, n.d.). For example, the gene symbols of the top upregulated genes (log FC >1.5) in NC were compared with the gene symbols of the top upregulated genes in NE (log FC >1.5).

Gene Ontology Enrichment Analysis - PANTHER

The overlapping gene dataset and the unique gene dataset for NC vs. NE differentiation were uploaded Gene Ontology Enrichment Analysis page to obtain molecular function analysis of significant genes in each dataset (Mi et al., 2019).
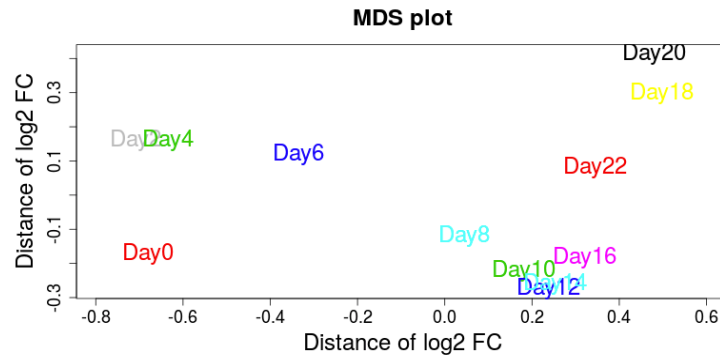
RESULTS

Fig. 1: Multidimensional scaling plot (MDS) for the neural ectoderm differentiation data (day 0 - 22) from Li et al. The MDS plot shows the variance between transcriptomic data obtained at different time points of NE differentiation. The variance between the datasets increases leading to day22. The plot shows how the data for each time point correlate with each other based on log fold change. The timepoints day 0, day2, day4, and day6 were used in our analysis.

Fig. 2: Volcano plot of differentially expressed genes from the compariosn of day 0 (ES cells) with day 2 prospective NE differentiation (Li et al.). The log 2 fold change values of the differentially expressed genes were plotted against the negative log 10 p-values. Parameter set for differential expression analysis: p-value (< 0.05) and fold change (>1.5 over ES)
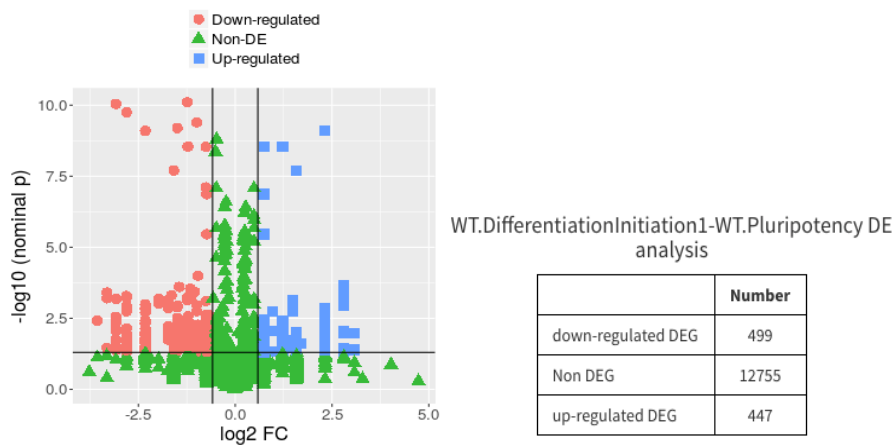
Fig. 3: Volcano plot of differentially expressed genes from the compariosn of day 0 (ES cells) with day 4 prospective NE differentiation (Li et al.). The log 2 fold change values of the differentially expressed genes were plotted against the negative log 10 p-values. Parameter set for differential expression analysis: p-value (< 0.05) and fold change (>1.5 over ES).
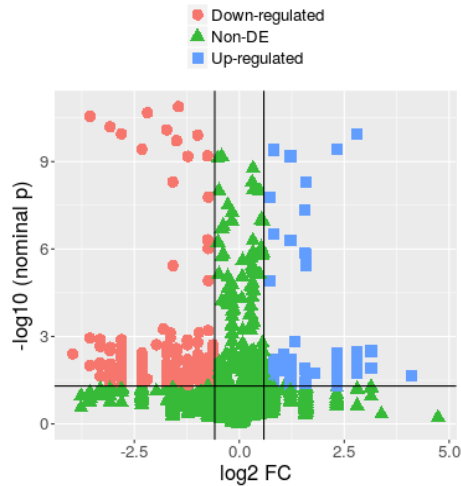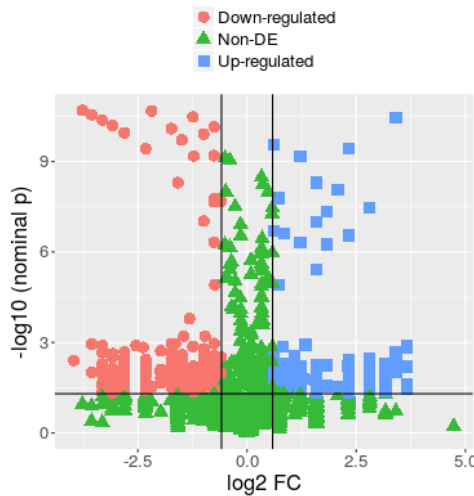

Fig. 4: Volcano plot of differentially expressed genes from the compariosn of day 0 (ES cells) with day 6 prospective NE differentiation (Li et al.). The log 2 fold change values of the differentially expressed genes were plotted against the negative log 10 p-values. Parameters set for differential expression analysis: p-value (< 0.05) and fold change (>1.5 over ES)
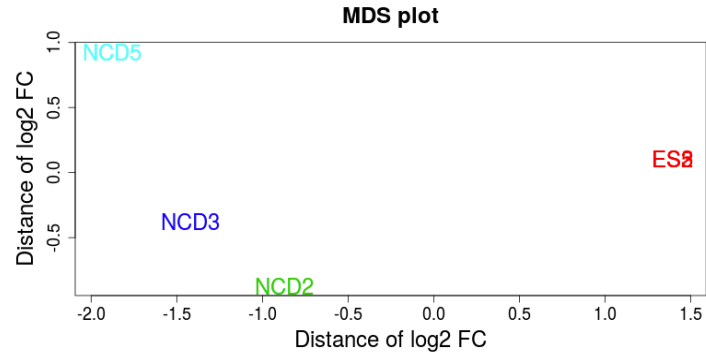
Fig. 5: Multidimensional scaling plot (MDS) for the neural ectoderm differentiation data from Prasad et al. The MDS plot shows the variance between transcriptomic data obtained at different time points of NC differentiation. The variance between the datasets increases leading to NCD5 (NC day 5). The timepoints ES (day 0), NCD2 (day2), and NCD3 (day3) were used in our analysis.
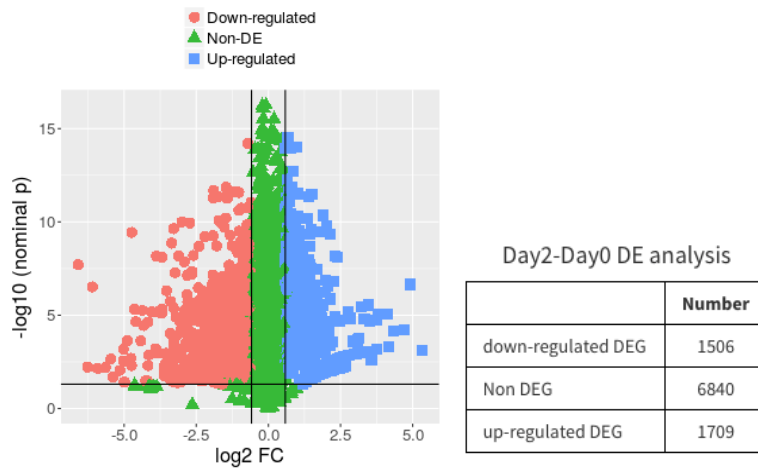
Fig. 6: Volcano plot of differentially expressed genes from the compariosn between day0 (ES cells) and day 2 prospective NC differentiation (Prasad et al.). The log 2 fold change values of the differentially expressed genes were plotted against the negative log 10 p-values. Parameters set for differential expression analysis: p-value ($< 0.05$) and fold change ($>1.5$ over ES). Similarly, gene expression analysis was carried out for day 4 and day 6 of prospective NE differentiation and the top gene hits data was obtained.
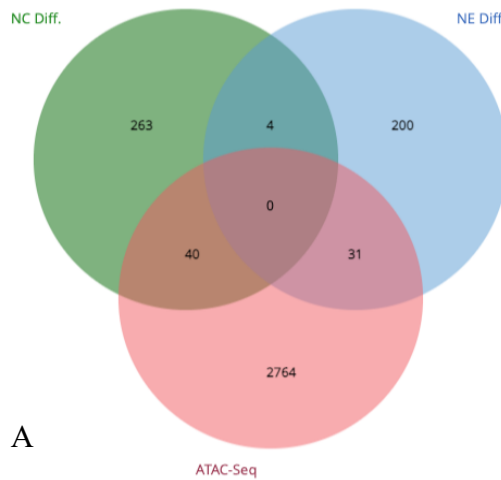
Fig. 7: Image A shows the comparison between **upregulated genes** in prospective NE differentiation day 2 (Li et. al) vs. prospective NC differentiation day 2 (Prasad et al.) vs. ATAC-seq data of chick embryos at early epiblast stage (HH4) (Williams et al.). Four genes appear to be common between the upregulated genes in NC differentiation day 2 vs. NE differentiation day 2. The top hits data from the limma voom differential expression analysis were futher filtered for comprattive analysis of the prospective NC differenation data and NE differnation data, the upregulated and down regulated genes were filtered by the parameters: log fold chang (>1.5) and P-value (<0.05).
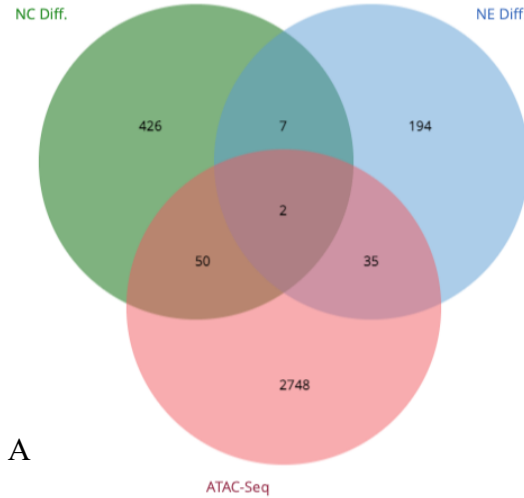
A

Fig. 8: The Venn diagram above shows the comparison between **downregulated genes** in prospective NE differentiation day 2 (Li et. al) vs. prospective NC differentiation day 2 (Prasad et al.) vs. ATAC-seq data of chick embryos at early epiblast stage (HH4) (Williams et al.). Two genes appear to be common between the three datasets. Seven of the genes appear to be common between NC differentiation day 2 and NE differentiation day 2. The top hits obtained from the limma voom differential expression analysis was further filtered for the comparative analysis. The parameter set to obtain the upregulated genes from the top hits dataset were log fold change (> 1.5) and p-value (<0.05).
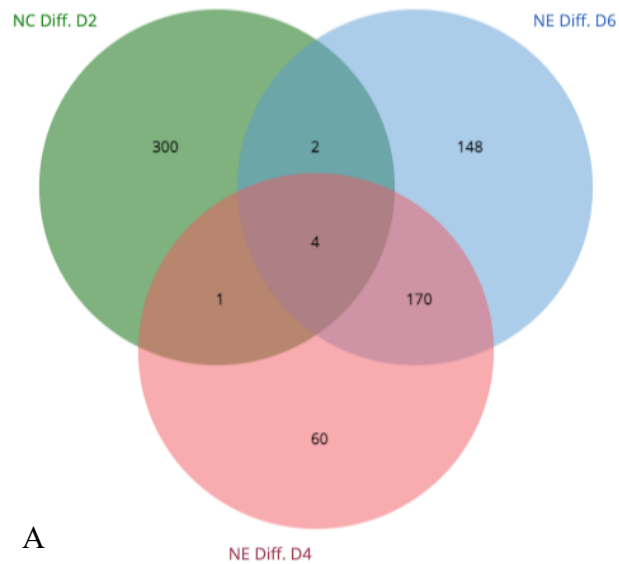
A

Fig. 9: The Venn diagram shows a comparison of **upregulated genes** in prospective NC differentiation day 2 (Prasad et al.) with prospective NE differentiation day 4 (Li et al.) and prospective NE differentiation day 6 (Li et al.). An overlap of four gene appears to be common among the three datasets. Two genes were common between NC differentiation day 2 and day 6. And one gene appears to be common between NC differentiation day 2 and 4. The top hits obtained from limma voom differential expression analysis were further filtered for the comparative analysis. The parameter set to obtain the upregulated genes from the top hits dataset were log fold change (> 1.5) and p-value (<0.05).

Table 1.1: The table below lists genes that overlap in the comparison of upregulated genes (log FC >1.5) in prospective NC and prospective NE differentiation day 2, and the genes unique to NC and NE differentiation at day2 (Fig. 7A).

Overlap in Upregulated Genes: NC vs. NE Differentiation (4 genes)	Top 25 Upregulated Genes Unique to NC Differentiation	Top 25 Upregulated Genes Unique to NE Differentiation	Top 25 Common Upregulated Genes in NC Differentiation and ATAC-Seq Data	Top 25 Common Upregulated Genes in NE Differentiation and ATAC-Seq Data
C5orf38	FST	GRHL3	GRB14	PDE3A
RMST	PKDCC	MIR3648	NRP2	FRZB
NR1I3	VAT1L	PLEKHB1	RGMB	MAF
SLC8A2	GRB14	SEMA6D	PLCL2	PAH
	UHRF1	CABP7	WWOX	TFAP2A
	RF02120	HHEX	EXOC6	MYB
	NRP2	PDE3A	SLC1A3	VIPR2
	RF02123	PPP1R1B	BMP2K	CDH8
	DACT1	EFNB1	GAP43	GABRE
	KCNN2	HTR1D	PACRGL	HHIP
	SOX21	KLF15	TRIM45	IGSF5
	TSPAN18	FRZB	ODC1	LYSMD2
	RGMB	MAF	IRX2	NLGN4Y
	RF00019-10	CCDC160	SACS	SLC24A3
	PLCL2	BNIPL	TMEM116	SLC38A4
	MALT1	BOC	GRAMD1B	SLC6A11
	CRNDE	PAH	CPS1	SMOC2
	AC0085221	TFAP2A	KCTD1	TRABD2B
	WWOX	TMEM35	CITED2	WDR72
	OSBPL11	BIK	TIAM1	ADAMTS3
	GBA	CACNG5	CCND1	CCDC171
	KIAA1586	CBLN1	CRIPT	CRYL1
	AC1068642	CHP2	OTULIN	FRRS1
	RF02180	DGAT2	TP53I11	HRH3
	CYP26A1	GABRA3	WDR35	SLC10A7

Table 1.2: The below table lists the genes that overlap in comparison of downregulated genes (log FC < -1.5) in prospective NC and prospective NE differentiation day 2, and the genes unique to NC and NE differentiation at day2 (Fig. 8A).

Overlap in Downregulated Genes: NC vs. NE Differentiation (7 genes)	Top 25 Downregulated Genes Unique to NC Differentiation	Top 25 Downregulated Genes Unique to NE Differentiation	Top 25 Common Downregulated Genes in NC Differentiation and ATAC-Seq Data	Top 25 Common Downregulated Genes in NE Differentiation and ATAC-Seq Data
MT1G	FOXD3-AS1	CAV2	EDNRB	ANKRD1
MT1E	MT1H	SNORA7B	DUSP6	CDH13
MT1F	LINC00458	LINC00152	RBM47	TIMP3
DLL3	AC0221402	SNORA76	KIZ	TOX2
SLC16A3	AC1042571	COL22A1	VLDLR	ADAMTS5
TNNT2	ADM	GCNT4	LRRK1	COL8A1
TSPO	RAB17	IGFBP6	COBL	DNER
	TDGF1	MMP9	ICA1	GFRA1
	RNU6-1330P	NPPB	PLCB1	NAV3
	AC0648021	SYTL2	PRICKLE1	WNT9A
	PMAIP1	GLIPR1	SPATS2L	RUNX1
	ACTA1	IGFBP7	PLOD2	ADAMTS15
	MT1X	KIFC3	RAB20	CDKN2B
	SPTSSB	MFAP5	CORO2A	CPEB2
	AP0009432	OSMR	ADCY2	HERC3
	MT2A	AHNAK2	PKP2	NOX4
	MT-TM	ANTXR2	MAP7	PLXNA4
	LINC01405	FAM157B	AASS	RGS6
	MT-TC	FOSL1	GRID2	BHLHE40
	RF00569	HBEGF	EDIL3	COL12A1
	AK4	IL32	RASL11B	ANGPTL4
	AC0050621	MYL7	SORBS2	CCNJ
	RF00019-2	NABP1	VSNL1	CHSY3
	RF02038	PDGFRB	SLC2A14	CISH
	RF00019-3	PLAT	SPRY2	GJD2
		GLIPR1		

Table 1.3: The below table lists genes that overlap in comparison of upregulated genes (log FC >1.5) in prospective NC differentiation day 3 vs. prospective NE differentiation day 4 vs. prospective NE differentiation day 6. The genes that are uniquely upregulated in prospective NC differentiation day 3, NE differentiation day 4, and NE differentiation day 6 are also listed (Fig. 9A).

Overlap in Downregulated Genes: NC day 2 vs. NE day 4 vs. NE day6 (6 genes)	Overlap in Upregulated Genes: NC day 2 vs. NE day 6 (4 genes)	Overlap in Upregulated Genes: NC day 2 vs. NE day 4 (1 gene)	Top 25 Upregulated Genes Unique to NC Differentiation day 2	Top 25 Upregulated Genes Unique to NE Differentiation day 4	Top 25 Upregulated Genes Unique to NE Differentiation day 6	Top 25 Upregulated genes common between NE Day 6 and NE day 4 Differentiation
SOX21	PLCL2	C5orf38	FST	SNORA81	GREB1L	DLK1
SLC8A2	NR1I3		PKDCC	DNAJC28	LHX5	EPHA4
RMST			VAT1L	CABP7	SIX3	FRZB
SHISA2			GRB14	CBLN1	LIX1	FEZF1-AS1
			UHRF1	SCNN1G	MAPK10	ZNF521
			RF02120	ACOT11	LHX2	HAPLN1
			NRP2	ADRA1A	PAX6	BOC
			RF02123	ALPPL2	RAX	CCDC160
			DACT1	ANKRD34B	SMOC1	DLL1
			KCNN2	ATP6V0A4	ZEB2	EFNB1
			TSPAN18	C10orf95	LRRC4B	GAS1
			RGMB	C9orf117	SOX5	LINC00571
			RF00019-10	CCM2L	TSPAN11	MAF
			MALT1	CFTR	VWF	PLEKHB1
			CRNDE	COX6B2	BCL6B	SCUBE3
			AC0085221	DPP4	BTBD17	SEMA6D
			WWOX	EPN3	BZRAP1	SKIDA1
			OSBPL11	FAM65C	C17orf96	SP8
			GBA	FAM78A	CCDC74B-AS1	TFAP2A
			KIAA1586	FUT2	CDH7	ARSG
			AC1068642	GABRE	CNTFR	BTG2
			RF02180	GFRA3	EFEMP1	C1R
			CYP26A1	HEXA	FAXDC2	CCDC15
			TEDC1	HHIP	FBLN7	CDH8
			USP3-AS1	HIST1H3J	FOXB1	CDH9

DISCUSSION

Our results for the comparison between prospective NE differentiation day 2 and prospective NC differentiation day 2 show that there is a limited overlap of genes that are expressed in the early stage of prospective neural crest differentiation and early stage of prospective ectoderm differential. The limited overlap may suggest that NC cells generated from pluripotent stem cells do not arise from neuroectodermal precursors. We identified four transcripts that overlapped in the comparison of upregulated genes in day 2 prospective NE and NC differentiation (Table 1.1). The four overlapping transcripts did not appear to code for important transcription factors therefore, they may not be playing important roles in lineage specification. From a Gene ontology enrichment analysis, important genes like TFAP2A, MAF, MYB, GHRL3, and KLF15 coding for transcription factors were found to be upregulated uniquely in prospective NE differentiation day 2.

Research carried out in chick cells proposes that the NC state is specified during the blastula stage from epiblast cells prior to fate segregation. It was identified that TFAP2A was expressed in an intermediate epiblast stage in the blastula epiblast, before gastrulation (Prasad et al., 2020). TFAP2A is also suggested to be a pioneer factor that is involved in the remodeling of the epigenomic landscape of progenitor cells to promote neural crest specification (Rothstein & Simoes-Costa, 2019).

A Gene Ontology enrichment analysis performed on the top upregulated genes unique to prospective NC differentiation day 2 revealed important genes coding for binding proteins, adaptor proteins, ion channels, and inhibitors. The genes NRP2 and GRB14 coding for important binding proteins were identified to be upregulated uniquely at prospective NC differentiation day 2. Knockdown studies on the genes coding for NRP receptor proteins have identified that they

are essential in NC cell migration, and the segregation of sensory and sympathetic fates to appropriate locations (Lumb et al., 2014). The gene KCNN2 coding for voltage-gated ion channels and CLCN2 coding for ion channels were also found to be uniquely in NC cells.

The SOX21 gene coding for an important transcription factor was identified to be uniquely upregulated in prospective NC differentiation day 2. A study of the SOX21 transcription factor in frog embryos suggests that high expressions of SOX21 play an inhibiting role in neuron formation and is known to promote progenitor maintenance (Whittington et al., 2015). The IRX2 gene coding for an important transcription factor responsible for neural specification was also upregulated (Freese et al., 2014)

In comparison to early chick epiblast ATAC-Seq data, we found unique overlaps pertaining uniquely to upregulated genes in prospective NC differentiation and prospective NE differentiation (Fig. 7B). The ATAC-Seq data indicates epigenetic factors critical for the expression of prospective NE and NC genes do not overlap. However, it cannot be inferred whether the open chromatin regions are actively transcribing DNA since ATAC-Seq only assesses if the chromatin regions are accessible or closed. To an extent, this supports the assumption that that critical transcript for NE is not observed in NC differentiation, suggesting that NC does not arise from ectoderm precursors.

We identified seven common gene transcripts that were downregulated in the comparison between NC and NE differentiation at day 2 (Table 1.2). The overlapping gene transcripts were not identified to code for important transcription factors. The gene DUSP6 coding for a protein phosphatase was found to be uniquely downregulated in the prospective NC cell differentiation. DUSP6 is known to act as a negative feedback regulator of FGF (fibroblast growth factor) and mutations in the gene can potentially cause FGFR related syndromes (C. Li et al., 2007). This

gene may play a crucial role in regulating FGF signaling that is responsible for the migration of NC during development (X.-T. Zhang et al., 2018)

In our comparison of the top upregulated genes in prospective NC differentiation day 2 with the upregulated genes in prospective NE differentiation day 4 and 6, we found critical transcripts that play major roles in lineage specification of the ectoderm (Table 1.3). We found important gene transcripts coding for essential transcription factors in neuroectoderm commitment. On running a Gene Ontology enrichment analysis on the upregulated genes, an important set of transcription factors were revealed. the transcripts ZNF521, SOX5, FOXB1, GAS1, EPHA4, SIX3, LHX2, LHX5, ZEB2, and GLIS3 coding for important transcription factors in NE specification were found to be upregulated at day 6 prospective NE differentiation. The neuroectoderm marker, PAX6 was also upregulated in the prospective NC day 6 differentiation. PAX6 is known to have a crucial downstream effect on neural inhibitors, and it is identified to be necessary for human neuroectoderm specification (Zhang et al., 2010). Studies have also confirmed that SIX3 plays an important role in neural differentiation. The knockout of SIX3 inhibited early neural differentiation and it was found to promote neural differentiation by regulating downstream transcription factors (Yuanyuan Li et al., 2017).

To recapitulate, our study supports the idea that NC cells are generated from pluripotent stem cells that do not emerge from neuroectodermal precursors. The limited gene overlap in the comparison between prospective NC differentiation day 2 and prospective NE differentiation day 2 suggests that NC cells may be endowed with the capacity to differentiate into mesodermal derivatives at early stages of gastrulation. If NC were to arise from neuroectodermal precursors, critical genes for neuroectoderm development should be observed in NC development. However, we did not observe a retention of NE or mesodermal gene expression in NC precursors.

The results for the comparison of prospective NC differentiation day 2 with NE differentiation suggested that at day 6 of prospective NE differentiation, the pluripotent cells are endowed with the potential to differentiation into NE-specific cells when important transcription factors governing the process of NE lineage specification are expressed.

A major limitation of this study is that the expression levels of gene sets may not be directly comparable with each other. Even though two genes appear to overlap with each other, the genes may have different expression levels in each data set. A WGCNA weighted co-expression network analysis will help homogenize or balance the expression levels to compare the genes directly with each other. Due to time constraints, a weighted WGCNA was not performed in our study.

The findings of this study were solely dependent on statistical modeling of gene expression data. All statistical methods are dependent on assumptions and they may fail to capture the biological significance behind data sets, and therefore, an accurate quantification will be hard to achieve.

Despite these limitations, our study supports previous findings on the expression of important transcripts and markers that are specified in prospective NE differentiation and NE differentiation. Early differentiation for prospective NE and NC revealed that NC do not arise form NE precursors. We produced a rich amount of data that can be further explored to identify important transcripts controlling the early differentiation of NC and NE cells. Our study further opens doors to explore how NC derived at the early gastrulation stage acquire and retain a broader differentiation potential.

REFERENCES

Dashboard—Genomics Biotools. (n.d.). Retrieved February 4, 2021, from

https://www.biotools.fr/

Freese, N. H., Lam, B. A., Staton, M., Scott, A., & Chapman, S. C. (2014). A Novel Gain-Of-

Function Mutation of the Proneural IRX1 and IRX2 Genes Disrupts Axis Elongation in the

Araucana Rumpless Chicken. PLoS ONE, 9(11).

https://doi.org/10.1371/journal.pone.0112364

Ladewig, J., Koch, P., & Brüstle, O. (2013). Leveling Waddington: The emergence of direct

programming and the loss of cell fate hierarchies. Nature Reviews Molecular Cell Biology,

14(4), 225–236. https://doi.org/10.1038/nrm3543

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). voom: Precision weights unlock linear

model analysis tools for RNA-seq read counts. Genome Biology, 15(2), R29.

https://doi.org/10.1186/gb-2014-15-2-r29

Leung, A. W., Murdoch, B., Salem, A. F., Prasad, M. S., Gomez, G. A., & García-Castro, M. I.

(2016). WNT/β-catenin signaling mediates human neural crest induction via a pre-neural

border intermediate. Development (Cambridge, England), 143(3), 398–410.

https://doi.org/10.1242/dev.130849

Li, C., Scott, D. A., Hatch, E., Tian, X., & Mansour, S. L. (2007). Dusp6(Mkp3) is a negative

feedback regulator of FGF stimulated ERK signaling during mouse development.

Development (Cambridge, England), 134(1), 167–176. https://doi.org/10.1242/dev.02701

Li, Y., & Andrade, J. (2017). DEApp: An interactive web interface for differential expression

analysis of next generation sequence data. Source Code for Biology and Medicine, 12.

https://doi.org/10.1186/s13029-017-0063-4

Li, Y., Wang, R., Qiao, N., Peng, G., Zhang, K., Tang, K., Han, J.-D. J., & Jing, N. (2017).

    Transcriptome analysis reveals determinant stages controlling human embryonic stem cell

    commitment to neuronal cells. The Journal of Biological Chemistry, 292(48), 19590–19604.

    https://doi.org/10.1074/jbc.M117.796383

Lumb, R., Wiszniak, S., Kabbara, S., Scherer, M., Harvey, N., & Schwarz, Q. (2014).

    Neuropilins define distinct populations of neural crest cells. *Neural Development*, *9*.

    https://doi.org/10.1186/1749-8104-9-24

Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2019). PANTHER version 14:

    More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools.

    Nucleic Acids Research, 47(D1), D419–D426. https://doi.org/10.1093/nar/gky1038

Prasad, M. S., Charney, R. M., & García-Castro, M. I. (2019). Specification and formation of the

    neural crest: Perspectives on lineage segregation. Genesis (New York, N.Y. : 2000), 57(1),

    e23276. https://doi.org/10.1002/dvg.23276

Prasad, M. S., Charney, R. M., Patel, L. J., & García-Castro, M. I. (2020). Distinct molecular

    profile and restricted stem cell potential defines the prospective human cranial neural crest

    from embryonic stem cell state. Stem Cell Research, 49, 102086.

    https://doi.org/10.1016/j.scr.2020.102086

Prasad, M. S., Uribe-Querol, E., Marquez, J., Vadasz, S., Yardley, N., Shelar, P. B., Charney, R.

    M., & García-Castro, M. I. (2020). Blastula stage specification of avian neural crest.

    Developmental Biology, 458(1), 64–74. https://doi.org/10.1016/j.ydbio.2019.10.007

Qin, D. (2019). Next-generation sequencing and its clinical application. Cancer Biology &

    Medicine, 16(1), 4–10. https://doi.org/10.20892/j.issn.2095-3941.2018.0055

Rothstein, M., & Simoes-Costa, M. (2019). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. Genome Research, gr.249680.119. https://doi.org/10.1101/gr.249680.119

Venn Diagrams—Genomics Biotools. (n.d.). Retrieved February 4, 2021, from https://www.biotools.fr/misc/venny

Waddington, C. H. (1942). Canalization of Development and the Inheritance of Acquired Characters. Nature, 150(3811), 563–565. https://doi.org/10.1038/150563a0

Whittington, N., Cunningham, D., Le, K., De Maria, D., & Silva, E. M. (2015). Sox21 regulates the progression of neuronal differentiation in a dose-dependent manner. Developmental Biology, 397(2), 237–247. https://doi.org/10.1016/j.ydbio.2014.11.012

Williams, R. M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U., Ling, I. T. C., Telenius, J., Taylor, S., Hughes, J., & Sauka-Spengler, T. (2019). Reconstruction of the Global Neural Crest Gene Regulatory Network In Vivo. Developmental Cell, 51(2), 255-276.e7. https://doi.org/10.1016/j.devcel.2019.10.003

Zhang, X.-T., Wang, G., Li, Y., Chuai, M., Lee, K. K. H., & Yang, X. (2018). Role of FGF signalling in neural crest cell migration during early chick embryo development. Zygote (Cambridge, England), 26(6), 457–464. https://doi.org/10.1017/S096719941800045X

Zhang, X., Huang, C. T., Chen, J., Pankratz, M. T., Xi, J., Li, J., Yang, Y., LaVaute, T. M., Li, X.-J., Ayala, M., Bondarenko, G. I., Du, Z.-W., Jin, Y., Golos, T. G., & Zhang, S.-C. (2010). Pax6 is a human neuroectoderm cell fate determinant. Cell Stem Cell, 7(1), 90–100. https://doi.org/10.1016/j.stem.2010.04.017