Video Article

RNA-Seq Analysis of Differential Gene Expression in Electroporated Chick Embryonic Spinal Cord

Felipe M. Vieceli¹, C.Y. Irene Yan¹

¹Department of Cell and Developmental Biology, Universidade de São Paulo

Correspondence to: Felipe M. Vieceli at fmvieceli@usp.br

URL: https://www.jove.com/video/51951

DOI: doi:10.3791/51951

Keywords: Developmental Biology, Issue 93, chicken embryo, in ovo electroporation, spinal cord, RNA-Seq, transcriptome profiling, Galaxy

workflow

Date Published: 11/1/2014

Citation: Vieceli, F.M., Yan, C.I. RNA-Seq Analysis of Differential Gene Expression in Electroporated Chick Embryonic Spinal Cord. J. Vis. Exp. (93),

e51951, doi:10.3791/51951 (2014).

Abstract

In ovo electroporation of the chick neural tube is a fast and inexpensive method for identification of gene function during neural development. Genome wide analysis of differentially expressed transcripts after such an experimental manipulation has the potential to uncover an almost complete picture of the downstream effects caused by the transfected construct. This work describes a simple method for comparing transcriptomes from samples of transfected embryonic spinal cords comprising all steps between electroporation and identification of differentially expressed transcripts. The first stage consists of guidelines for electroporation and instructions for dissection of transfected spinal cord halves from HH23 embryos in ribonuclease-free environment and extraction of high-quality RNA samples suitable for transcriptome sequencing. The next stage is that of bioinformatic analysis with general guidelines for filtering and comparison of RNA-Seq datasets in the Galaxy public server, which eliminates the need of a local computational structure for small to medium scale experiments. The representative results show that the dissection methods generate high quality RNA samples and that the transcriptomes obtained from two control samples are essentially the same, an important requirement for detection of differential expression genes in experimental samples. Furthermore, one example is provided where experimental overexpression of a DNA construct can be visually verified after comparison with control samples. The application of this method may be a powerful tool to facilitate new discoveries on the function of neural factors involved in spinal cord early development.

Video Link

The video component of this article can be found at https://www.jove.com/video/51951/

Introduction

Genetic studies on live organisms frequently use the chicken embryo as a model because *in ovo* electroporation represents a fast and inexpensive way to partially transfect embryonic structures *in vivo* with DNA constructs¹⁻⁵. Bicistronic expression vectors that encode one transcript containing a fluorescent reporter together with the gene of interest, such as pCIG⁶ or pMES⁷, allow quick verification of transfection quality in the desired region under a stereomicroscope before processing embryos for downstream analysis.

With recent advances in DNA sequencing, it is now possible to obtain digital whole transcriptome expression profiles from RNA samples using RNA-Seq^{8,9}. Therefore, instead of time-consuming methods that enable analysis of only a few gene products in parallel, such as *in situ* hybridization, immunohistochemistry and qPCR, preparation of cDNA libraries for high-throughput sequencing can provide information of the whole transcriptome. Observation of the experimental effects on the expression levels of every single gene may in turn provide powerful insights on the pathways altered by the construct used. This is particularly easy if a genomic assembly for the organism used is available, thus the chick embryo again is a suitable model.

If the user has access to a reliable genome sequencing center, the main issue is obtaining high quality RNA samples. This work demonstrates a method for obtaining RNA samples from transfected neural tubes suitable for RNA-Seq, as well as the downstream steps for *in silico* analysis of the resulting datasets. After electroporation at HH12-13 followed by 48 hr incubation, transfected trunk spinal cord halves were dissected in ribonuclease-free conditions, immediately lysed and further protected from degradation by ultra-low freezing until RNA purification and library preparation.

The Galaxy public server¹⁰ provides access to free resources for analyzing high-throughput sequencing data. The amount of space offered for registered users is enough for small experiments, thus eliminating the need for local computational infrastructure. RNA-Seq data analysis includes filtering, alignment and quantitation/comparison of gene expression. Although there are general guidelines for RNA-Seq data analysis, filtering parameters will largely depend on the quality of sequencing and should be determined after quality control analysis of the raw data.

This protocol describes the steps to obtain and compare RNA-Seq profiles from chicken embryonic spinal cords transfected *in vivo*. As a representative result, comparison of datasets obtained from two independent control samples transfected with an empty vector showed that the method is reproducible and should allow identification of differentially expressed transcripts in experimental samples. Therefore, application of



this method has the potential to greatly increase the number of discoveries after one single experiment and thus provide a large set of data for subsequent investigation.

Protocol

1. Electroporate the Construct of Interest in the Neural Tubes of HH12-13 Chicken Embryos In ovo

Use a fluorescent reporter, preferably in a bicistronic transcript or fused to the protein of interest with an intercalating viral 2a peptide¹¹, to enable quick identification of individuals with satisfactory levels of transfection. NOTE: Typical incubation time for this stage is around 48 hr at 37.8 °C.

- Open a small window in the eggs after removing 3 ml of albumin with a syringe coupled to a 18 G x 1 1/2 needle. To enhance embryo
 visualization, inject a small amount of Indian ink diluted 1:50 in sterile Ringer's solution underneath the embryo¹⁻⁴.
- After covering the embryo with sterile Ringer's solution, inject a 10 mM Tris-HCl, pH 8.2 and 0.1% Food Dye (F, D & C) solution containing 2.5 μg/μl DNA through the posterior end of the neural tube until it fills up to the hindbrain region. To electroporate, position platinum electrodes (0.5 mm in diameter) separated by 4 mm¹ along the flank region and deliver 5 pulses of 50 msec and 20 V with an interval of 100 msec between them⁴.
- 3. After electroporation, incubate manipulated embryos for 48 hr more until they reach stage HH23 and identify individuals presenting high level of transfection at the flank region under a fluorescent stereomicroscope.
 NOTE: The duration of incubation after electroporation depends on the aim of the experiment. The electroporated plasmids can generate significant expression levels after 2 hr². Thus, if the interest lies in early onset genes the post-electroporation incubation period can be reduced.

2. Harvest Successfully Transfected Embryos and Dissect the Electroporated Spinal Cord Halves

NOTE: This procedure lasts from 2-3 hr for each 10 embryos and yields 1-2 µg total RNA per embryo. The dissection of neural tubes requires fine movements and some practice sessions may be necessary before applying it to experimental samples.

- Decontaminate metallic dissection tools with RNase decontamination solution or by baking at 150 °C for 4 hr¹² and prepare DEPC-treated PBS. Use only new certified RNase-free plasticware. NOTE: After harvesting, it is important to follow standard care to prevent RNase contamination during dissection.
- 2. Harvest the embryos using a pair of fine dissecting scissors and a sieved collection spoon. Keep the embryos in cold DEPC-treated PBS in a
- 3. With a pair of fine tweezers, cut out the whole flank region (between the posterior limit of the anterior limb bud and the anterior limit of the posterior limb bud) and transfer this section to a new dish containing cold DEPC-treated PBS.
- 4. Use two fine glass needles to perforate at three different points on each dorsolateral side of the embryo between the neural tube and the
- 5. Introduce the two glass needles in the central perforation and slowly move one of them away along the neural tube anterior-posterior axis. Repeat this for the perforations that were not reached by the first movements. Remove the detached paraxial mesoderm with a pair of fine tweezers.
- 6. Turn the embryo sideways, insert the tips of the tweezers between the neural tube and the notochord and slide them in opposite directions along the anterior-posterior axis to separate the two structures.
 - NOTE: This should also remove most of the mesoderm still attached to the neural tube.
- 7. If necessary, remove any larger pieces of mesodermic tissue that remain attached by pulling with a pair of tweezers while lightly holding the neural tube with another pair of tweezers. When finished, gently transfer the neural tube to another dish containing cold DEPC-treated PBS using a RNase-free glass Pasteur pipette coupled to a manual pipette pump and keep on ice until all the tubes are dissected.
- 8. Divide the neural tubes into two halves. For this, insert the sharp tip of a closed pair of tweezers in the lumen and open the roof plate by moving the tweezers out through it. After this, separate the halves by cutting the floor plate with the tweezers tips.
- 9. Sort electroporated from non-electroporated halves under a fluorescent dissecting scope and transfer to a 1.5 ml tube filled with ice-cold DEPC-treated PBS. If necessary, flick the microtubes gently to release any tissue attached to the plastic walls and to promote sedimentation of the tubes.
- 10. Spin at 100 x g for 3 sec and remove very slowly as much PBS as possible with a 1 ml micropipette. Do this while looking at a light source that shines through the microtube to monitor if any embryonic material is sucked to the pipette tip; if this happens, slowly discharge and repeat the centrifugation.
- 11. Add 70 µl of lysis solution for RNA extraction per neural tube half and flick the tube to mix contents. Incubate at room temperature for 5 min and vortex vigorously. Repeat the 5 min incubation and vortexing two more times. Store at -80 °C.
- 12. Collect at least eight neural tube halves per sample to ensure an adequate yield for the next step. Pool smaller samples if several electroporation/dissection sessions are required to amass sufficient number of neural tubes. Store sufficient samples for duplicates or triplicates for each condition.

3. Purify RNA and Prepare Libraries for Illumina Sequencing

- 1. Remove samples from -80 °C storage and place on ice for 15-30 min. Resume thawing at room temperature until complete.
- 2. Vortex well and leave at room temperature for 5 min. Vortex again and centrifuge at 15,000 x g for 1 min to pellet insoluble material.



- 3. Transfer the supernatant to a new tube and proceed to RNA extraction according to the total RNA isolation kit manual instructions. Elute in the minimum recommended volume to increase concentration. NOTE: DNase treatment after clean-up is recommended in order to avoid potential contamination of genomic DNA sequences in RNA-Seq datasets.
- 4. Flick the tube to mix contents and remove 5 µl for quality control analysis. Store the remaining sample at -80 °C until library preparation; these samples should be thawed only once to avoid RNA degradation.
- 5. Run purified RNA in a Bioanalyzer instrument using the RNA 6000 Nano kit; RIN value must be at least above 7 to proceed.
- Generate libraries using an Illumina compatible RNA-Seq library preparation kit and multiplex up to 4 libraries in one lane of a HiSeq instrument set to generate 100 bp single reads. If possible, include all samples in the same run to minimize potential technical artifacts. NOTE: This should yield up to 50 million reads per sample, which is more than enough to obtain coverage for most genes¹³.

4. Compare RNA-Seq Datasets in the Galaxy Public Server¹⁰.

NOTE: Server storage space is limited and experiments comparing many samples may require deletion of intermediate files generated in the filtering steps.

- Create a user account in the Galaxy public server (https://usegalaxy.org) and FTP upload the Illumina datasets in fastq format. Prepare uploaded files with the FastqGroomer 14 tool under the NGS:QC and manipulation menu; if the quality scores are already encoded in Phred+33 (Sanger), skip this step by simply changing the file type to fastqsanger under the Attributes for each dataset.
- Analyze the general quality of each dataset using the FastQC tool under the NGS:QC and manipulation menu. NOTE: Special attention should be given to the resulting graphics for per base sequence quality and per base sequence content, as these will provide estimates for the parameters used for trimming the ends of the reads.
- Trim reads to discard 3' ends with low quality data using the tool FASTQ Quality Trimmer¹⁴ under the NGS:QC and manipulation menu. Set the tool to process only 3' ends for quality score below 20.
- Remove adapter sequences from reads using the Clip tool (http://hannonlab.cshl.edu/fastx_toolkit/) under the NGS:QC and manipulation menu. Set the tool to discard reads with less than 45 bp left, input the sequence of the adapter used and choose the output to include both clipped and non-clipped sequences. Additionally, opt not to discard sequences with unknown bases since most low quality bases were already removed in the last step.
- Using the tool Trim sequences (http://hannonlab.cshl.edu/fastx_toolkit/) under the NGS:QC and manipulation menu, trim the first bases of the reads to remove 5' sequence bias detected by FastQC (typically 10). NOTE: At this point, it is a good idea to run FastQC again and compare with the previous analysis to determine if the dataset quality was
- 7. Download and unpack the latest chicken genome assembly for the UCSC genome browser from Illumina iGenomes (http:// support.illumina.com/sequencing/sequencing_software/igenome.ilmn) and upload the whole genome fasta format file and the genes annotation gtf format file to the Galaxy history using the tool Upload File under the Get Data menu.
- Map reads for each sample to the chicken genome using the aligner TopHat2¹⁵ under the NGS: RNA Analysis menu. Choose to use a genome from History (select the fasta file uploaded in **step 4.7**). Build a predicted transcriptome for each sample using *Cufflinks*¹⁶ under the *NGS: RNA Analysis* menu. Choose to use reference annotation
- as a guide (select the gtf file uploaded in step 4.7) and turn on the options Perform Bias Correction (from History, fasta file from step 4.7) and Use multi-read correct.
- 10. Merge predicted transcriptomes for all samples using Cuffmerge 16 under the NGS: RNA Analysis menu. Choose to use reference annotation (gtf file from step 4.7) and sequence data (from History, fasta file from step 4.7).
- 11. Compare BAM files generated by TopHat2 for each sample using Cuffdiff¹⁶ under the NGS: RNA Analysis menu. In the field Transcripts select the merged transcriptome generated by Cuffmerge and turn on the options Perform Bias Correction (from History, fasta file from step 4.7) and Use multi-read correct.
- 12. Sort differential expression testing tables numerically ascending in column 13 (q value) using the tool Sort under the Filter and Sort menu. NOTE: Genes/transcripts showing significant differential expression between experimental conditions will present the lowest q values.
- 13. To check results visually with coverage graphs, first convert BAM files to bedGraph format using the tool Create a BedGraph of genome coverage 17 under the BEDTools menu. Disable reporting of regions without coverage, choose to treat spliced entries as distinct intervals and scale the coverage by a factor that normalizes sequencing depth for all the datasets being compared.
- 14. Convert the resulting bedGraph file to bigWig format using the tool Wig/BedGraph-to-bigWig (http://www.genome.ucsc.edu/) under the Convert Formats menu. Upload the bigwig files as custom tracks in the UCSC genome browser (http://www.genome.ucsc.edu/) for the galGal4 assembly and search a gene of interest to visualize the coverage in the datasets.

Representative Results

To validate the method described, two control samples were generated from embryos electroporated with the empty vector pMES⁷ and one sample from embryos transfected with the same vector containing the insert construct SCRT2-ZNF, which encodes the zinc-finger domains of chicken SCRT2¹⁸. Samples yielding 11-22 µg total RNA each were obtained from pools of 8 to 12 embryos. All three samples scored an optimal RIN value of 10 in a quality check performed with the Bioanalyzer instrument (Figure 1), demonstrating that high quality RNA can be obtained with this protocol.

After sequencing, the described filtering process increased the alignment rate from 83 to 91% of reads (data not shown). As expected, comparison of the transcriptomes from two control samples showed no clear difference when globally compared in a scatter plot (Figure 2) and the Pearson correlation coefficient was 0.99. Additionally, display of coverage for all the three datasets along the locus that contains the portion of the SCRT2 sequence showed a clear increase in the number of reads mapped to this region in the experimental sample (Figure 3), indicating that the overexpression was successful and that differences between control and experimental samples should be detectable with the method described here.

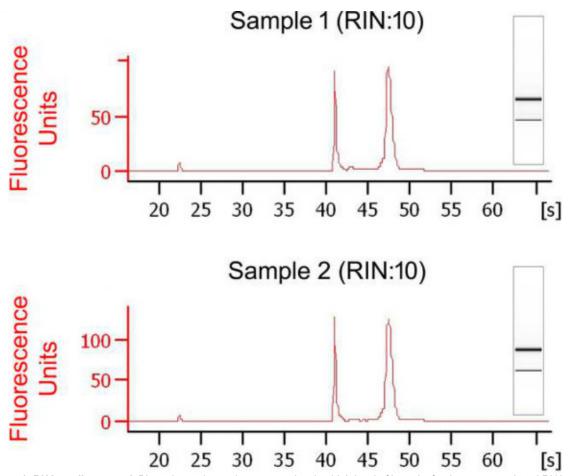


Figure 1: RNA quality control. Bioanalyzer electropherograms showing high level of integrity for the two control total RNA samples indicated by the 18S and 28S ribosomal RNA peaks. Insets represent a virtual gel for the run.

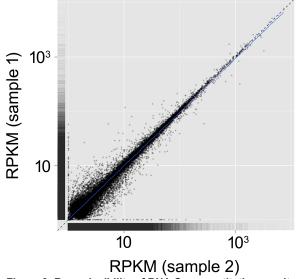


Figure 2: Reproducibility of RNA-Seq quantitation results. Scatter plot showing that RPKM for most genes is the same in both control samples. Virtually all genes that do not follow this pattern have very low RPKM values and thus are usually ignored from transcriptome comparisons.

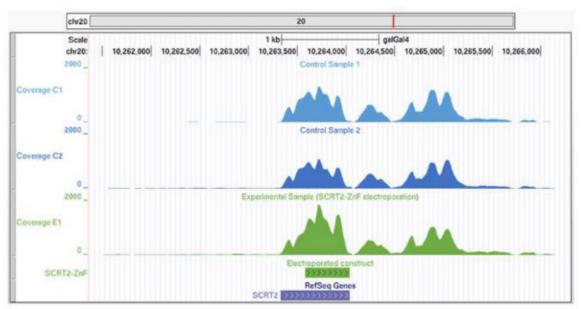


Figure 3: Visualization of differences in an experimental sample. UCSC Genome Browser display of the SCRT2 locus in the chicken genome (galGal4) showing a visible increase in coverage of the region that codes for the protein domain (SCRT2-ZnF) ectopically expressed in the experimental sample.

Discussion

Here we provide guidelines for analyzing effects after electroporation of the chicken spinal cord. Although electroporation of DNA vectors is more frequently used to overexpress a gene of interest, one can also use constructs encoding dominant negatives, quimeric proteins or precursors for siRNAs to generate gene function knock-down conditions¹⁹⁻²¹. In fact, comparison of transcriptome profiles resulting from both gain- and loss-of-function analysis may point out genes differentially expressed in opposite ways in each condition, thus having the potential to uncover more specific effects.

As any method applied downstream to RNA extraction, synthesis of RNA-Seq libraries requires a minimum initial amount of starting material. With this protocol, 8 embryos should be sufficient to yield around 10 µg of total RNA. Starting with a lower number of individuals without compromising an adequate working concentration is probably possible with the use of RNA extraction systems that allow final dilution in smaller volumes.

Proper segregation of the neural tube from all mesoderm should be practiced a few times before applying to important samples, as this is a delicate step that requires some familiarity. To avoid excessive RNA degradation during the dissection step, it is also important to complete the whole procedure for all neural tubes in about 1-2 hr. Again, practice will be important to improve dissection time. One important observation is that the connection between mesoderm tissues is stronger than the connection between mesoderm and neural tube, thus pulling mesodermal structures, specially the notochord, is more efficient than trying to cut unwanted tissue. Also, researchers should feel free to develop their own technique optimized to their own dissection conditions. However, care should be taken for samples compared in a same experiment to be dissected similarly so that level of contamination with mesodermal tissues is as homogeneous as possible between samples.

Our protocol demonstrates the comparison of transcriptome profiles from stage HH23 chicken embryonic spinal cords transfected at stage HH12-13. However, small adjustments should allow its application to embryos harvested at earlier stages. On the other hand, if later stages are desired for analysis of the end result, the bicistronic vectors mentioned may not be suitable due to a continuous reduction in the number of molecules per cell as development proceeds. Therefore, especially if the intended end point is achieved after more than 72 hr after transfection, it may be necessary to use vectors encoding transposon mediated genomic integration of the construct being used ²².

Another issue to be considered is the heterogeneity of the electroporation efficiency. The number of non-transfected cells ranges from a few to a considerable amount, especially the ones in the ventral neural tube. Since after dissection all cells are subjected to downstream analysis, a low electroporation efficiency will reduce the assay's sensitivity. An approach to enrich the sample with electroporated cells is the use of fluorescence activated cell sorting (FACS) to reduce the population of non-transfected cells from the final sample²³. If such an approach is used, however, a greater number of embryos will be necessary to achieve sufficient yield for RNA-Seq library preparation.

Illumina sequencing is recommended as this is more widely offered by sequencing centers and *in silico* RNA-Seq data analysis is simplified when compared to other available platforms²⁴. However, other options are still adequate with modifications in data filtering steps more suitable to the platform chosen. RNA-Seq data analysis guidelines are still not completely established and comparison of outputs from different available tools, some of which are already included in the Galaxy public server, is highly recommended. Nevertheless, the *in silico* analysis described here should be suitable for obtaining good comparison results from high quality datasets. As in all other areas in Biology, high-throughput sequencing increases enormously the amount of information that can be obtained from a single experiment, and the method described here is one way of applying it to the study of spinal cord development.

Disclosures

The authors have nothing to disclose.

Acknowledgements

CYIY is supported by FAPESP (2012/14421-5) and FMV is supported by a fellowship from FAPESP (2009/53695-0). We thank the DNA Technologies Core at University of California, Davis for preparation and sequencing of the RNA-Seq libraries used in this work, the Galaxy team for providing an excellent and free interface for high-throughput sequencing data analysis tools and Dr. Marianne Bronner for allowing us to film in her laboratory space.

References

- 1. Itasaki, N., Bel-Vialar, S., Krumlauf, R. 'Shocking' developments in chick embryology: electroporation and in ovo gene expression. *Nat. Cell Biol.* **1** (8), E203 207, (1999).
- 2. Krull, C.E. A primer on using in ovo electroporation to analyze gene function. Dev. Dyn. 229 (3), 433 439, (2004).
- 3. Blank, M.C., Chizhikov, V., Millen, K.J. In ovo electroporations of HH stage 10 chicken embryos. *J. Vis. Exp.* (9), 408, doi: 10.3791/408, (2007).
- 4. Sauka-Spengler, T., Barembaum, M. Gain- and loss-of-function approaches in the chick embryo. *Methods Cell Biol.* 87, 237 256, doi: 10.1016/S0091-679X(08)00212-4, (2008).
- Ghadge, G.D., et al. Truncated wild-type SOD1 and FALS-linked mutant SOD1 cause neural cell death in the chick embryo spinal cord. Neurobiol. Dis. 21 (1), 194 – 205, (2006).
- Megason, S.G., McMahon, A.P. A mitogen gradient of dorsal midline Wnts organizes growth in the CNS. Development. 129 (9), 2087 2098, (2002).
- 7. Swartz, M., Eberhart, J., Mastick, G.S., Krull, C.E. Sparking new frontiers: using in vivo electroporation for genetic manipulations. *Dev. Biol.* 233 (1), 13 21, (2001).
- 8. Cloonan, N., et al. Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat. Methods. 5 (7), 613 619, doi: 10.1038/nmeth.1223, (2008).
- 9. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods.* **5** (7), 621 628, doi: 10.1038/nmeth.1226, (2008).
- 10. Goecks, J., Nekrutenko, A., Taylor, J., Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11** (8), R86, doi: 10.1186/gb-2010-11-8-r86, (2010).
- 11. Trichas, G., Begbie, J., Srinivas, S. Use of the viral 2A peptide for bicistronic expression in transgenic mice. *BMC Biol.* **6**, 40, doi: 10.1186/1741-7007-6-40. (2008).
- 12. Maniatis, T., Fritsch, E.F., Sambrook, J. *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1982).
- 13. Wang, Y., et al. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. BMC Bioinformatics. 12, 10:S5, doi: 10.1186/1471-2105-12-S10-S5, (2011).
- 14. Blankenberg, D., et al. Manipulation of FASTQ data with Galaxy. Bioinformatics. 26 (14), 1783 1785, doi: 10.1093/bioinformatics/btq281, (2010).
- 15. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14** (4), R36, (2013).
- Trapnell, C., et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 28 (5), 511 – 515, doi: 10.1038/nbt.1621, (2010).
- 17. Quinlan, A.R., Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26** (6), 841 842, doi: 10.1093/bioinformatics/btq033, (2010).
- 18. Vieceli, F.M., Simões-Costa, M., Turri, J.A., Kanno, T., Bronner, M., Yan, C.Y. The transcription factor chicken Scratch2 is expressed in a subset of early postmitotic neural progenitors. *Gene Expr. Patterns.* **13** (5-6), 189 196, doi: 10.1016/j.gep.2013.03.004, (2013).
- 19. Chesnutt, C., Niswander, L. Plasmid-based short-hairpin RNA interference in the chicken embryo. Genesis. 39 (2), 73 78, (2004).
- 20. Katahira, T., Nakamura, H. Gene silencing in chick embryos with a vector-based small interfering RNA system. *Dev. Growth. Differ.* **45** (4), 361 367, (2003).
- 21. Rao, M., Baraban, J.H., Rajaii, F., Sockanathan, S. In vivo comparative study of RNAi methodologies by in ovo electroporation in the chick embryo. *Dev. Dyn.* **231** (3), 592 600, (2004).
- 22. Lu, Y., Lin, C., Wang, X. PiggyBac transgenic strategies in the developing chicken spinal cord. *Nucleic Acids Res.* **37** (21), e141, doi: 10.1093/nar/gkp686, (2009).
- 23. Simões-Costa, M., Tan-Cabugao, J., Antoshechkin, I., Sauka-Spengler, T., Bronner, M.E. Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. *Genome Res.* **24** (2), 281 290, doi: 10.1101/gr.161182.113, (2014).
- 24. Trapnell, C., et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7 (3), 562 578, doi: 10.1038/nprot.2012.016, (2012).