Musical Genes: A Shiny Application to listen to the transcriptional symphony of parental care in pigeons

02 Feb 2021
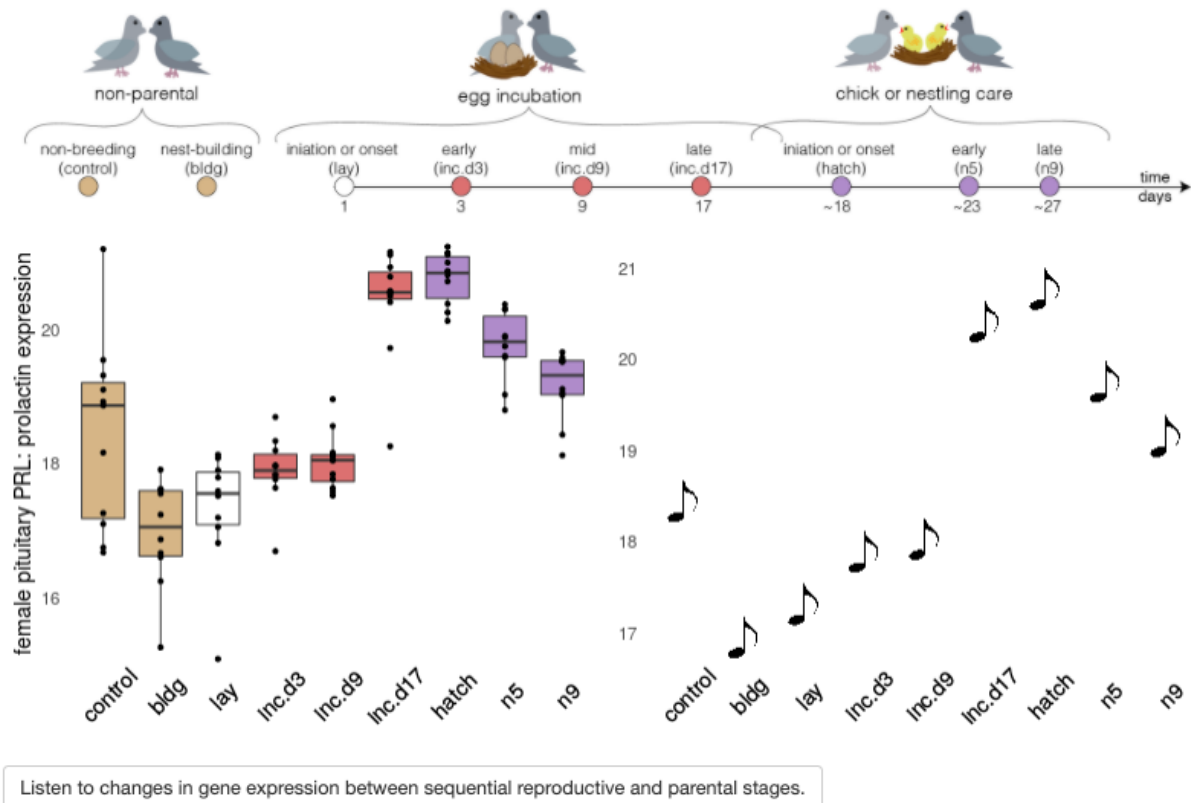
Summary

In music, a "symphony" is an elaborate instrumental compositions written for an orchestra. In biology, the word symphony is used as a metaphor to describe the harmonious (or cacophonous) combination of elements that orchestrate cellular activities (e.g. 2020 Lemon Tijan). With RNA-sequencing, it is possible to quantify the expression of genes thought to regulate complex biological processes, but most RNA-seq workflows are designed to report results as figures and tables (rather than sound) and require bioinformatic expertise. To fill this gap, we built an R Shiny Application that allows interactive data exploration, visualization and sonification the transcriptomic symphony regulating parental care in pigeons. With a few clicks of a button, user can interactively examine the raw data for genes of interest, listen to how the gene changes over time and relation to a few key reproductive hormones. This tools makes the research more accessible to scientists, peer-reviewers, the general public, and to musicians who asipre to make data-inspire music.

Approach

This application uses transcriptomic data from Austin & Harris et al 2021 and hormone data from Austin et al 2021. These studies use a highly replicated experimental design to characterize the transcriptional activity of the hypothalamus, pituitary, and gonads axis at nine time points over the course of reproduction and parental care in male and female pairs as they transitioned from a sexually mature, non-reproductive state to a reproductive one. Nine pseudo-timepoints measured represent nest building, egg incubation, nesting care. Data were imported into R Studio and stored in duckdb for faster access. Data were joined, filtered, selected, mutated, and plotted with various R packages (). The mean value of gene expression was calculated for each group and sonified those values using the R packages `sonify` and `tuneR` (ref). (Fig 1). R Shiny was used to build and deploy an interactive web application. (https://raynamharris.shinyapps.io/musicalgenes/). All the R code available under CC-BY at https://github.com/raynamharris/musicalgenes.

Transcriptional symphony

Listen to changes in gene expression between sequential reproductive and parental stages.

The sound works well when the shiny app is deployed locally, but it does not work on all browsers. To overcome this limiartion, we calculated and rounded the mean value of gene expression for each time point and and converted the value to the letters A to G so that the user can play a representation of the data on an instrument on their choosing, such as a piano (Fig 2). In the future, we would like to play the sounds of many genes at once by a digital or live orchestra to listen to coordinated or discordiated changes in gene expression.

gene	instrument	notes
AR: androgen receptor	violin	AGEEECCEF
AVP: arginine vasopressin	viola	GADAABABB
AVPR1A: arginine vasopressin receptor 1A	saxophone	BAEGEDABA
CRHBP: corticotropin releasing hormone binding protein	keyboard	GDFECABEC
ESR1: estrogen receptor 1	upright bass	AAGDCBCCA
FOS: Fos proto-oncogene, AP-1 transcription factor subunit	french horn	GACEDDBAB
GALR1: galanin receptor 1	clarinet	CECECEFGA
GNRHR: gonadotropin releasing hormone receptor	tuba	GGAEDDBCF
OXT: oxytocin/neurophysin I prepropeptide	cello	GABABABAA
PGR: progesterone receptor	trombone	EDACBAACG
PRL: prolactin	piano	CABBCFGED
PRLR: prolactin receptor	oboe	GAABBCBBA
TH: tyrosine hydroxylase	bassoon	BGBCBBABC
VIP: vasoactive intestinal peptide	trumpet	DCDGABCDE

Next we sought to explore the phsiological symphony that results from the interplay between genes and hormones. Austin et al. 2021 in press, measured cirulating prolactin, sex steroids (estradiol and testosterone), corticosterone, and progesterone and found week correlations between prolactin and corticosterone. With this Shiny application, we can explore correlations between hormones and gene expression (Fig 3). To listen the physiological symphony, we sonified the hormone concentration to see if it increases, decreases, or doesn't change as the gene of interest increases in expression.

Listen to changes in hormone concentration as gene expression increases.
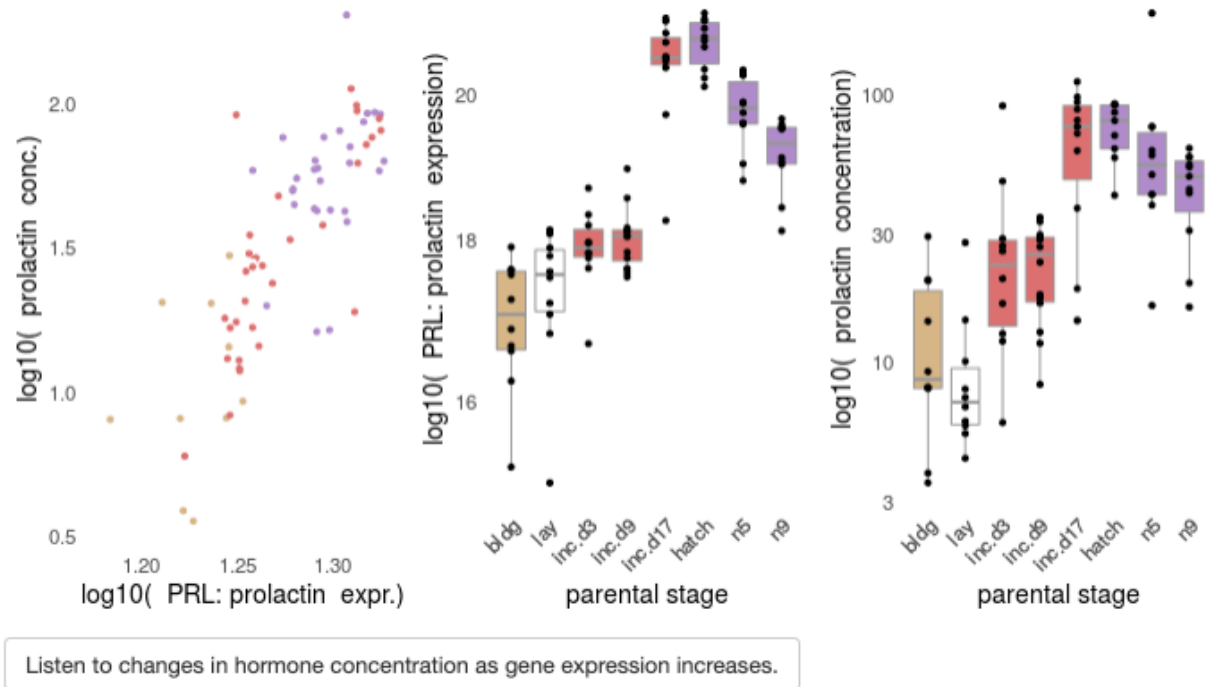
Because many users may not be familiar with many of the genes in this dataset, we important gene names for the human orthologs from the HUGO database (ref). Finally, we added additional information about genes, such as their description, ontology, and associated diseases were obtained from the gene ontology database (Genome alliance).

Gene description

Predicted to have hormone activity and prolactin receptor binding activity. Involved in several processes, including negative regulation of endothelial cell proliferation; positive regulation of pri-miRNA transcription by RNA polymerase II; and positive regulation of receptor signaling pathway via JAK-STAT. Predicted to localize to extracellular space and secretory granule. Implicated in carotid artery disease. Biomarker of hypoglycemia; obesity; and severe acute respiratory syndrome.

Gene Ontology: Associated Biological Processes

female pregnancy; lactation; mammary gland development; maternal behavior; negative regulation of endothelial cell proliferation; positive regulation of cell population proliferation; positive regulation of epithelial cell proliferation; positive regulation of lactation; positive regulation of pri-miRNA transcription by RNA polymerase II; positive regulation of receptor signaling pathway via JAK-STAT; regulation of ossification; regulation of receptor signaling pathway via JAK-STAT; response to nutrient levels; signal transduction

Associated diseases

carotid artery disease; hyperandrogenism; hyperprolactinemia; hypoglycemia; obesity; portal hypertension; prolactinoma; renovascular hypertension; severe acute respiratory syndrome

One limiation of this application is that is is directly tied to this particular data sature. A future implementation of the tools could allow flexible data input. For instance, BioJupies is an online tool that rapidly creates interactive visualization of any RNA-seq data set in NCBI or a user uploaded data set (Torre et al). Our software applies to a single data set. The complexity of our experimental design with multiple variables and many combinations of comparison, scaling the data wrangling and visualization is not feasible. Regardless of size, future studies benefit from tools like BioJupies or Musical Genes or rapid data assessment. Despite these limitations, we still believe that this application presents a unique contribution to science by providing open access to software that makes data and results more accessible to diverse audiences.

In summary, we have created an application that facilitates data exploration and hypothesis testing for scientist and clinicians that can also inspire artists to create music to better understand symphonies in a biolgoical context. This application is interactive, so it allows users to explore the raw data in a way that is not possible when reading a published manuscript saved as a pdf. By simultaneously viewing the data, reading the statistical summary, and listening the the changes in mean expression, the user gains a deeper understand at the complexity of the data and the biology and brings us one step closer to listening to the transcriptional symphony of parental care.

References

Citations to entries in paper.bib should be in rMarkdown format.

If you want to cite a software repository URL (e.g. something on GitHub without a preferred citation) then you can do it with the example BibTeX entry below for (**???**).

For a quick reference, the following citation commands can be used: - `@author:2001` -> "Author et al. (2001)" - `[@author:2001]` -> "(Author et al., 2001)" - `[@author1:2001; @author2:2001]` -> "(Author1 et al., 2001; Author2 et al., 2002)"

Acknowledgements

We acknowledge contributions from x and y.

References