GENOME-WIDE IDENTIFICATION AND VALIDATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN THE LEFT AND RIGHT OVARIES OF THE EMBRYONIC CHICKEN

BY

MOHD SAJEED
ID No. RVM/2020-20

THESIS SUBMITTED TO

P.V. NARSIMHA RAO TELANGANA VETERINARY UNIVERSITY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE OF
MASTER OF VETERINARY SCIENCE
(ANIMAL GENETICS AND BREEDING)
IN THE FACULTY OF VETERINARY SCIENCE

DEPARTMENT OF ANIMAL GENETICS AND BREEDING
COLLEGE OF VETERINARY SCIENCE
P.V. NARSIMHA RAO TELANGANA VETERINARY UNIVERSITY.
RAJENDRANAGAR, HYDERABAD - 500 030

January, 2023

CERTIFICATE

This is to certify that **Mr. MOHD SAJEED (RVM/20-20)** has satisfactorily prosecuted the course of research and that the thesis entitled **"GENOME-WIDE IDENTIFICATION AND VALIDATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN THE LEFT AND RIGHT OVARIES OF THE EMBRYONIC CHICKEN"** submitted is the result of original research work done and is of a sufficiently high standard to warrant its presentation to the examination. I also certify that the thesis or part thereof has not been previously submitted by her for a degree of any University.

Date: 27-1-23

Place: Hyderabad

(Dr. B SRIDEVI)

Chairman of Advisory Committee

CERTIFICATE

This is to certify that the thesis entitled "**GENOME-WIDE IDENTIFICATION AND VALIDATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN THE LEFT AND RIGHT OVARIES OF THE EMBRYONIC CHICKEN**" submitted in partial fulfilment of the requirements for the degree of **Master of Veterinary Science** of **P. V. Narsimha Rao Telangana Veterinary University, Hyderabad** is a record of *bonafide* research work carried out by **MOHD SAJEED (RVM/2020-20),** under our guidance and supervision. The subject of the thesis has been approved by the Student's Advisory Committee.

No part of the thesis has been submitted by the student for any other degree or diploma. The published part has been fully acknowledged. All assistance and help received during the course of investigations have been duly acknowledged by the author.

Thesis approved by the Student Advisory Committee

Dr. B. SRIDEVI Chairperson .....
Associate Professor,
Department of Animal Genetics and Breeding,
College of Veterinary Science, Korutla,
Hyderabad, 500030.

Dr. P. JAYA LAXMI Member .....
Professor and Head
Department of Animal Genetics and Breeding,
College of Veterinary Science,
Mamnoor, Warangal, 506166.

Dr. P. KALYANI Member .....
Assistant Professor,
Department of Animal Biotechnology,
College of Veterinary Science, Rajendranagar,
Hyderabad, 500030.

Dr. S. JAYAKUMAR Member .....
Senior Scientist (Animal Genetics & Breeding)
ICAR - Directorate of Poultry Research,
Rajendranagar, Hyderabad, 500030.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

&	:	And
%	:	Percent
µM	:	Micro moles
>	:	Greater than
±	:	Plus or Minus
°C	:	Degree Celsius
µg	:	Microgram
µl	:	Microlitre
AMH	:	Anti Mullerian Hormone
APOC3	:	Apolipoprotein C-III
APOH	:	Apolipoprotein H
AHSG	:	Alpha 2-HS Glycoprotein
AR	:	Androgen Receptors
ASZ1	:	Ankyrin Repeat, SAM And Basic Leucine Zipper Domain Containing 1
BAHS	:	Basic Animal Husbandry Statistics
BMP7	:	Bone Morphogenetic Protein 7
BMP3	:	Bone Morphogenetic Protein 3
bp	:	Base Pair
cDNA	:	Complementary DNA
CGS	:	Candidate gene sequencing
circRNAs	:	Circular RNA
CPB2	:	Carboxypeptidase B2
C_T value	:	Cycle threshold value
CYP19A1	:	Cytochrome P450 family 19 subfamily A member 1
DAZL	:	Deleted in AZoospermia
DDX4	:	DEAD-Box Helicase 4
DDBJ	:	DNA Data Bank of Japan
DEPC	:	Diethyl pyrocarbonate
DMRT1	:	Doublesex And Mab-3 Related Transcription Factor 1
DNA	:	Deoxyribo Nucleic acid

DEGs	:	Differentially expressed genes
dNTP	:	Deoxyribonucleotide triphosphate
dsDNA	:	Double-stranded deoxyribonucleic acid
D1	:	Day 1
EBI	:	European Bioinformatics Institute
E6	:	Embryonic 6th Day
E8	:	Embryonic 8th Day
E10	:	Embryonic 10th Day
E12	:	Embryonic 12th Day
E18	:	Embryonic 18th Day
EMB	:	Embigin
ERNI	:	Early response to neural induction
ESR1	:	Estrogen receptor 1
FAO	:	Food and Agriculture Organization
FET1	:	Female-specific expression transcription factor
FABP1	:	Fatty Acid Binding Protein 1
FETUB	:	Fetuin B
FGA	:	Fibrinogen alpha chain
FGB	:	Fibrinogen beta chain
FGG	:	Fibrinogen gamma chain
FSH	:	Follicle-stimulating hormone
FST	:	Follistatin
GAL	:	Galanin and GMAP Prepropeptide
GAPDH	:	Glyceraldehyde-3-phosphate dehydrogenase
GBS	:	Genotyping by sequencing
GDF-8	:	Myostatin
GFF	:	General feature formats
GO	:	Gene Ontology
GREM1	:	Gremlin 1
GTF	:	Gene transfer format
h	:	Hour
HH	:	Hamburger and Hamilton’s
HISAT2	:	Herarchical indexing for spliced alignment of transcripts
HKGs	:	Housekeeping genes

HMGCS2	:	3-Hydroxy-3-Methylglutaryl-CoA Synthase 2
HNF-3	:	Hepatocyte nuclear factor-3
HRG	:	Histidine rich glycoprotein
HSI	:	Hepatosomatic index
HTS	:	High-throughput sequencing
ICAR	:	Indian Council of Agricultural Research
IL-8	:	Interleukin -8
IGF1	:	Insulin-like growth factor 1
ITIH2	:	Inter-alpha-trypsin inhibitor heavy chain 2
lncRNAs	:	Long noncoding RNAs
INSDC	:	International Nucleotide Sequence Database Collaboration
KDa	:	Kilodalton
KEGG	:	Kyoto Encyclopedia of Genes and Genomes
KNG1	:	Kininogen 1
LEAP2	:	Liver Enriched Antimicrobial Peptide 2
miRNA	:	MicroRNAs
min	:	Minutes
mm	:	Millimeter
mL	:	Milliliter
mRNA	:	Messenger RNA
MOV10L1	:	Mov10 Like RISC Complex RNA Helicase 1
NCBI	:	National Center for Biotechnology Information
NFW	:	Nuclease Free Water
NGS	:	Next-generation sequencing
NROB1	:	Nuclear Receptor Subfamily 0 Group B Member 1
NR5A2	:	Nuclear Receptor Subfamily 5 Group A Member 2
ng	:	Nano gram
nm	:	Nanometer
OD	:	Optical density
OVALY	:	Ovalbumin
PAX2	:	Paired box gene 2
PCR	:	Polymerase Chain Reaction
PGCs	:	Primordial germ cells

PPAR	:	Peroxisome proliferator-activated receptors
qPCR	:	Quantitative polymerase chain reaction
QTL	:	Quantitative trait loci
RNA	:	Ribonucleic acid
RNAi	:	RNA interference
RNA-seq	:	RNA sequencing
rRNA	:	Ribosomal ribonucleic acid
RPM	:	Revolutions per minute
RT-qPCR	:	Real-time quantitative PCR
RQ	:	Relative Quantitation
SERPINA10	:	Serpin family A member 10
SERPINB2	:	Serpin family B member 2
SERPINC1	:	Serpin family C member 1
SERPIND1	:	Serpin family D member 1
SHC4	:	SHC Adaptor Protein 4
SMC1B	:	Structural Maintenance of Chromosomes 1B
SOLiD	:	Sequencing by Oligo Ligation Detection
SOX	:	SRY-related HMG box
SOX9	:	SRY-Box Transcription Factor 9
SRA	:	Sequence Read Archive
SYCP3	:	Synaptonemal Complex Protein 3
Taq	:	Thermus aquaticus
TBE	:	Tris Boric acid EDTA
TDRD12	:	Tudor Domain Containing 12
TERB1	:	Telomere Repeat Binding Bouquet Formation Protein 1
TF	:	Transferrin / ovotransferrin
TEX14	:	Testis Expressed 14, Intercellular Bridge Forming Factor
tRNA	:	Transfer ribonucleic acid
TOM1L1	:	Target Of Myb1 Like 1 Membrane Trafficking Protein
TUBA3E	:	Tubulin Alpha-3E Chain
UTR	:	Untranslated region
WES	:	Whole exome sequencing
WNT4	:	Wingless-related MMTV integration site 4
β-actin	:	Beta actin

DECLARATION

I, **MOHD SAJEED (RVM/20-20)** hereby declare that the thesis entitled **"GENOME-WIDE IDENTIFICATION AND VALIDATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN THE LEFT AND RIGHT OVARIES OF THE EMBRYONIC CHICKEN"** submitted to P.V. Narsimha Rao Telangana Veterinary University for the degree of **MASTER OF VETERINARY SCIENCE** is a result of original research work done by me. It is further declared that the thesis or any part thereof has not been submitted for any other degree or diploma.

Date: 27/01/2023

Place: Hyderabad

(MOHD SAJEED)

ACKNOWLEDGEMENTS

It is a great pleasure to look back and recall the path one traverses through, during those trying times and day's perseverance. It is indeed a privilege at this juncture to recall all the faces and spirits in the form of teachers, friends, near and dear ones. Above all, I praise the Almighty Lord for his abundant grace, mercy and blessings showered on me in each and every step of my life.

I express my gratitude and sincere thanks to the chairman of my advisory committee ***Dr. B. Sridevi,*** *Associate Professor, Department of Animal Genetics and breeding, College of Veterinary Science, Korutla, for her mentorship and encouragement to complete my Master's course work and thesis. She was always there to listen and to give advice throughout the course and research. I found it a privilege to work under her. I sincerely thank her for showing different ways to approach research problems and need to be persistent to accomplish any goal.*

I would like to express my sincere thanks to ***Dr. Jayakumar S****, Senior Scientist, Animal Genetics and Breeding, ICAR Directorate of Poultry Research, Rajendranagar, Hyderabad, and member of the advisory committee for his wholehearted support, for providing timely needed chemical and equipment's to carryout un-interrupted research, for his mentorship and encouragement to complete my Master's research, an ideal teacher, mentor, and research supervisor, offering advice and encouragement with a perfect blend of insight and knowledge. a tremendous spirit of working and thank you for helping me in this investigation.*

I would like to express my sincere thanks to ***Dr. P. Jaya Laxmi*** *, Professor and Head, Department of Animal Genetics and breeding, College of Veterinary Science, Mamnoor, and member of the advisory committee for her valuable suggestions, critical comments and assistance during the period of research work and in preparation of the synopsis and thesis.*

I express my gratitude to ***Dr. P. Kalyani****, Assistant Professor, Animal Biotechnology, College of Veterinary Science, Rajendranagar, Hyderabad, and member of the advisory committee for helping me in guidance with research work and also thanks for helping guidance with results, discussion and writing up the thesis.*

I express my sincere thanks to ***Dr. D. Sakaram****, Professor and Head, Department of Animal Genetics and Breeding, College of Veterinary Science, Korutla*

for his valuable suggestions, critical comments and assistance during the period of research work and in preparation and valuable corrections of thesis.

I express my sincere thanks to ***Dr. D. Sreenivas****, Professor and Head, Department of Animal Genetics and Breeding, College of Veterinary Science, Rajendranager for his valuable suggestions, critical comments and advisory for research.*

My sincere thanks to ***Dr. T K Bhattacharya****, National fellow, Animal Genetics and Breeding, ICAR Directorate on Poultry Research, Rajendranagar, Hyderabad for allowing me to work in his laboratory whenever necessary and providing the materials to work with.*

My sincere gratitude to ***Dr. M. Shanmugam,*** *Senior Scientist, Physiology, ICAR Directorate on Poultry Research, Rajendranagar, Hyderabad for his valuable suggestions, helping in embryonic tissue collections with great patience and accuracy, encouragement during research period.*

My sincere gratitude to ***Dr. R.N. Chatterjee,*** *Director, ICAR Directorate on Poultry Research, Rajendranagar, Hyderabad for providing an opportunity for research and enhancing my skills at ICAR Directorate on Poultry Research.*

Let me put on record my deep sense of gratitude to ***Dr. J. Sai Prasanna****,* ***Dr Swathi****,* ***Dr Dr.K. Sarin Kunnath*** *, staff of Deptartment of Animal Genetics and Breeding, College of Veterinary Science, Rajendranaga for their teachings & help during my research work.*

I owe my deepest gratitude to my seniors ***Dr. B. Rajith Reddy, Dr. Parthasarathi, Dr. Thejasri Jakkoju, Dr. Pragna, Dr Shivani, Dr Nagasri, Dr Azmeera Suvarna, Dr Vasanthi, Dr. K Ramesh****, for their continuous moral support, providing a homely environment and technical knowledge, valuable suggestions, assistance and encouragement throughout my post-graduation life and manuscript preparation, without whose support I could not complete this piece of work in time.*

I use this opportunity to sincerely thank my colleagues ***Dr. P. Sankeerthi, Dr. Lalitha Shree, Dr. Vasanthi, Dr. Tejaswi, Dr. Ragini, Dr. Tulasi, Dr. T Mounica, Dr. Suman, Dr. P Basha, Dr. D Naveen, Dr. Hanuman, Dr. T Venkatesh, Dr. Sharath Nayak, Dr. Ramraj Sagar, Dr. Hemanth, Dr. Boda Sai, Dr. Nihal Reddy, Dr. Chandu,***

Dr. Shyam babu, Dr. Rajkumar, Dr. Krishna Chitanya, *for their mentorship, friendship, love, help and care and for making the two-year study very much enjoyable and memorable.*

I am extremely delighted to extend my thanks to my junior colleagues ***Dr. Mullu Atchutrao, Dr Amarnath, Dr Vamshi, K Gyaneshwar Yadav*** *for their timely help rendered during the period of study and research.*

I am dearth of words to express my unbound gratitude and love to my beloved venerable parents ***Mohd Chand Pasha*** *and* ***Shaheen begum****, for providing me invaluable moral support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.*

I owe my deepest gratitude to my friends ***P. Rakesh, B, Vijay Bhaskar, Mohd Jani, Dr Bhanu Prakesh, G. Lokesh, Dr Mohd Anwar, Chandra Shekar, K Krishna Kanth Reddy, C Anil Kumar,*** *for shaping and tuning my career, for there eternal support and understanding of my goals and aspirations.*

I would like to thank Library members – ***Dr. Atchamamba,*** *Assistant Professor,* ***Rajesh, Kishore and Balamani*** *for their kind co-operation during my research work.*

I wish to express my thanks to ***Madhu and Kumar*** *non-teaching staff of Department of Animal genetics and breeding, College of Veterinary Science, Rajendranagar, Hyderabad and* ***Raju, Santosh*** *non-teaching staff ICAR Directorate on Poultry Research, Rajendranagar, Hyderabad*

I am thankful to ***P. V. Narsimha Rao Telangana Veterinary University, Hyderabad*** *for providing financial assistance in the form of a stipend during my postgraduate study.*

I wish to extend my thanks to one and all those who have contributed even in a small way either direct or indirect manner to bring out my work in the present form.

(MOHD SAJEED)

Name of the Author	:	**MOHD SAJEED**
Title of the Thesis	:	**GENOME-WIDE IDENTIFICATION AND VALIDATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS IN THE LEFT AND RIGHT OVARIES OF THE EMBRYONIC CHICKEN**
Degree to which it is submitted	:	**MASTERS OF VETERINARY SCIENCE**
Faculty	:	**VETERINARY SCIENCE**
Department	:	**ANIMAL GENETICS AND BREEDING**
Major Advisor	:	**Dr. B. SRIDEVI**
University	:	**P.V. NARSIMHA RAO TELANGANA VETERINARY UNIVERSITY**
Year of submission	:	**2022**

ABSTRACT

The present study was conducted to identify the differentially expressed genes in left and right chicken ovaries and their validation at ICAR-Directorate of Poultry Research, Rajendranagar, Hyderabad. In asymmetric chick gonads, the left and right female gonads go through different developmental processes, producing a functioning ovary exclusively on the left side. Despite considerable advancements in recent years, little is known about the genomics of the right ovary degeneration in chick embryo, and the mechanisms of molecular regulation remain poorly understood.

In this study, transcriptome analysis was used to look into the left and right ovaries differently expressed genes and the functions of the genes at $E12^{th}$ day. Transcriptome data of Embryonic 12^{th} day was utilised from NCBI SRA data bank. It was discovered that there were 72.7 million readings in all of the collected raw data. The quality of FASTQ files was checked. 53.06 million of the processed reads were uniquely mapped to the *Gallus gallus* reference genome (GRCg7b), yielding a mapping percentage of 80.22. In the current study, 9152 genes were revealed to be significantly differently expressed between the embryonic left and right chicken ovaries on $E12^{th}$ day. Similarly, the total number of significantly up-regulated genes by a fold change above +2 in the right ovary during E12 was 479 and the significantly down-regulated genes by fold change below -2 in the right ovary during E12 were 227.

Pathways such as Blood Clotting Cascade pathway, PPAR signalling pathway, Neuroactive ligand-receptor interaction, Insulin signalling pathway, Ferroptosis, Apoptosis, Nuclear division, Fibrinolysis, Endopeptidase inhibitory activity, Blood coagulation, Meiotic cell cycle process are involved with the differentially expressed genes between right and left ovaries causing right ovary regression.

For Validation of transcriptome analysis ten fertile eggs of 12^{th} day incubation were randomly chosen and were sacrificed. Four left and right ovaries and liver were collected from chicken embryos and stored in RNA later solution at -80^{o}C till further processing.

Total RNA was isolated using TRIzol reagent from both the ovaries. After assessing the concentration and purity of isolated RNA, cDNA was synthesized using the reverse transcription kit. Specific primers were utilized and the qPCR conditions were standardized for the two up and down regulated genes (FETUB, LEAP2, DAZL and DDX4). The relative expressions of these genes were normalized using the housekeeping gene (β-actin). In comparison to the left ovaries (24.57±1.95), the right ovaries (22.61±0.73) showed higher levels of FETUB gene expression with a lower C_T. Similar to this, right ovaries showed higher LEAP2 gene expression levels with lower C_T (21.59±0.46) as compared to left ovaries (22.58±2.13). In comparison to the left ovaries (18.53±1.36), the right ovaries (22.67±1.13) showed lower levels of DAZL gene expression with greater C_T. Similar to the left ovaries (19.433±1.73), right ovaries (22.63±0.68) showed lower levels of DDX4 gene expression with greater C_T. Right ovaries showed the highest expression of FETUB and LEAP2, whereas the lowest expression of DAZL and DDX4, in accordance with the transcriptome data validating it.

INTRODUCTION

CHAPTER - I

INTRODUCTION

The Chicken (*Gallus gallus domesticus*) is one of the most widespread domesticated poultry species. Among different cultures, religions, and societies across India. Chicken is widely accepted with little or no taboo. India's poultry population is around 851.81 million, which includes both commercial and backyard fowl. Commercial poultry has a population of 534.74 million, whereas backyard poultry has a population of 317.07 million (20th Livestock Census). Chicken serves as an important source of protein through meat and egg production around the world. India ranks sixth in the world in terms of meat output, having increased from 6.7 million tonnes in 2014-15 to 8.6 million tonnes in 2019-20, with a 5.98 % annual growth rate (FAO STAT, 2019). Chicken egg production has great economic importance. India accounts for around 7.22 percent of world egg production (FAO, 2021) and ranks 3rd in egg production (20th Livestock Census). Table eggs are abundant in nutrients and viable eggs are the backbone of the chicken business. The ovary is the primary site of egg production. The total egg and meat production from our country is around 114.38 billion eggs and 4.34 million tonnes respectively, with an annual commercial egg production being around 95.17 billion eggs. India ranks 5th in broiler meat production in the world (BAHS, 2020).

In chickens, during the gonadal differentiation, male birds develop symmetrical bilateral testes, whereas female birds develop only a left ovary. The chicken's right ovary is present embryologically, but it starts regressing during mid-embryonic development and is no longer functional in adult birds. Only the left ovary and oviduct are found in birds, which evolved as a flying adaption over years (Zheng *et al.*, 2013). This evolutionary adaptation decreases the aircraft's weight, which is required for flight. The majority of modern birds have only 1 ovary and oviduct, flying bird fossils from the Cretaceous period [125 million years ago (mya)] have a single ovary on the left side, while the non-flying dinosaur ancestors of birds had 2 oviducts and 2 ovaries (Zheng *et al.*, 2013). The evolutionary process of regression of the right ovary is incorporated into their genome. In addition to the morphological asymmetry between the left and right gonads, several genes are related to asymmetrical development and right ovary degeneration. In chickens, sex-specific differentiation of gonads is evident

from embryonic day 6–6.5 (E6–6.5/ stage 29–30) and there is an asymmetry in the development between left and right gonads after E8 (stage 34) (Smith, 2007). An unusual asymmetry in gonadal development is observed in most avian species including ducks (Van *et al.,* 1960) and chickens. The present study is to know the differentially expressed genes in the left and right ovary during embryonic development involved in the regression of the right ovary. In chickens, several genes are related to asymmetrical development and right ovary degeneration (Andrews *et al.,* 1997; Hoshino *et al.,* 2005; Guioli *et al.,* 2007). A deeper study through functional genomics can help to identify the genes regulating the right ovary regression.

Researchers have since focussed on a deeper understanding of the chicken genome to study the expression of transcripts (Shaikat *et al.,* 2018) involving ovary development and how they are regulated in different phases of the life cycle. Variation in gene expression has been identified as an important mechanism underlying susceptibility to complex diseases and evolutionary processes. The concept of functional genomics (Moore *et al.,* 2005) has evolved that focuses on understanding the function and regulation of genes and gene products on a genome-wide scale. The functional genomics approach can be used to simultaneously investigate differences in the expression of thousands of genes concerning environmental and physiological challenges. This approach allows us to determine which group of genes are responsible for metabolic/pathologic changes and how these genes may be manipulated to improve performance and productivity.

With the introduction of Next Generation Sequencing (NGS), bioinformatics and other high-throughput technologies, it is now possible to study the transcriptome(Lowe *et al.,* 2017), or the entire set of RNA transcripts produced by the genome, under specific conditions, in specific cells, or at different times or phases of cell growth and differentiation (Slatko *et al.,* 2018). A transcriptome captures a snapshot in time of the total transcripts present in a cell (Lowe *et al.,* 2017). Rapid improvements and widespread use of different high-throughput experimental methods have resulted in an abundance of omics data in various researches in recent years, and appropriate analytical tools are required to properly interpret the biological information contained within this data. These "-omics" data, which all end with the same suffix, are used to investigate an organism's genetic material ("Genomics"), RNA transcripts

("Transcriptomics"), proteins ("Proteomics"), epigenetic modification ("Epigenomics"). RNA sequencing (RNA-seq) and microarray are the two main technologies for quantifying the transcriptome. An organism's or a cell's total number of RNA molecules, or transcripts which includes messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and other non-coding RNAs like microRNA (miRNA) combinedly called transcriptome (Ding *et al.*, 2021). The main class of RNA molecules known as "mRNA" is responsible for transmission of genetic information from DNA into proteins. The transcriptome reflects genes that are expressed at a particular moment and can change depending on many environmental factors, in contrast to the genome, which is essentially fixed for all the cells in their life time.

The conventional Sanger sequencing has been now replaced by next-generation sequencing (NGS) (ER Mardis - Annu Rev Anal and 2013 [no date]) for omic data analysis during the last few years. The advent of NGS technology has changed how we think about scientific methodologies in fundamental, applied, and clinical research. NGS can analyze millions of sequence reads simultaneously and has been extensively used for genome sequencing (including whole genome and whole exome sequencing), transcriptome profiling (RNA-seq), DNA methylation (bisulfite sequencing), DNA-protein interaction (ChIP-sequencing), etc.

In a typical RNA seq (Marguerat and Bähler 2010) workflow, the whole RNA is converted into cDNA, which are chopped into fragments with specific oligos adapter (referred to as reads) then sequenced using suitable platform or technology(Attia and Saeed 2016) (e.g. Illumina, SOLiD, Roche). The generated RNA-seq data is processed through different bioinformatic tools for differential gene expression analysis. Genes that are differentially expressed in various cell types or at different stages of the life cycle can be identified by comparing transcriptomes. During various stages of cell development and differentiation, altered gene expression profiles may reveal common patterns of response. This approach allows us to determine which groups of genes are responsible for asymmetric ovarian development and right ovary regression.

Objectives:

1. Identification of differentially expressed transcripts in the left and right ovaries during the chicken embryonic development.

2. To validate the identified transcripts involved in ovarian development.

CHAPTER - II

REVIEW OF LITERATURE

The domestic chicken (*Gallus gallus domesticus*), the most widely spread of poultry, supplies people with consistent amounts of protein in the form of meat and eggs. In Chicken, during embryonic development, several genes encoding transcription factors and secreted growth factors play a crucial role in left and right ovary asymmetry. Presently, 37.2% of all poultry in India are native chickens and these birds produce about 17.8% of all eggs produced in the nation (Kanakachari *et al.,* 2022).

2.1 CHICKEN GONAD DEVELOPMENT

The overall embryogenesis time in chickens is 21 days. The gonads form throughout development as thickenings along the ventromedial surface of the mesonephric kidneys (developmental stage 22; day 3.5) (Hamburger and Hamilton 1992). The gonads of both sexes consist of a germinal epithelial layer comprising somatic and germ cells, as well as an underlying medulla of epithelial cords interspersed with mesenchymal cells, before sexual differentiation. By stage 30, the commencement of gonadal sex differentiation may be recognized histologically (day 6.5).

Although both the left and right ovaries are present in the embryonic stages of all birds, only the left ovary develops fully into a functional ovary. It has been noted that both the left and right gonads are typically functioning in falconiformes and brown kiwis. (Johnson 2015). A rare case of a persistent and functional right ovary in hens and ducks was evident (Sturkie and Mueller 1976). The following wild bird species were found to have a right ovary: long-eared owl (*Asio otus*), common buzzard (*Buteo buteo*), sparrow hawk (*Accipiter nisus*), and goshawk (*Accipiter gentilis*) (Rodler *et al.,* 2015).

Functional right oviduct was discovered in various chicken breeds (Champion 1955) and (Sell 1959). Chicken embryos gained the maximum weight throughout the embryogenesis during 12-13 days of incubation. During this phase, it was noticed the greatest growth in ovarian mass in the chicken embryo (Khokhlov *et al.,* 2020).

2.1.1 Male Gonad Development

The male bird has a homomorphic (ZZ) chromosome, whereas the female is heteromorphic (ZW). The right and left testes to become tubular structures about 3 mm long and 0.5 mm wide and while the right female gonad is similar in appearance but slightly smaller than the testes in males, the left female gonad has acquired a broader, flatter appearance and is significantly larger at about 3.2 mm long and 0.8 mm wide (Guioli *et al.*, 2014). However, DMRT1 is essential for chicken testis determination, and the Z dose hypothesis for avian sex determination is validated(Smith *et al.*, 2009).

2.1.2 Female Gonad Development

In females, only the left functional ovary and oviduct are developed and the right ovary and oviduct become regressed. The differences in the right and left ovary and oviduct starts at the embryonic stage itself. In females during embryonic development, the right ovary is around 2.6 mm by 0.5 mm, while the left ovary, is approximately 8 mm by 1.5 mm (Guioli *et al.*, 2014). The majority of modern flying birds have only one ovary and one oviduct, and bird fossils from the Cretaceous period [125 million years ago (mya)] have a single ovary on the left side, whereas non-flying dinosaur ancestors of birds had two oviducts and two ovaries (Zheng *et al.*, 2013).

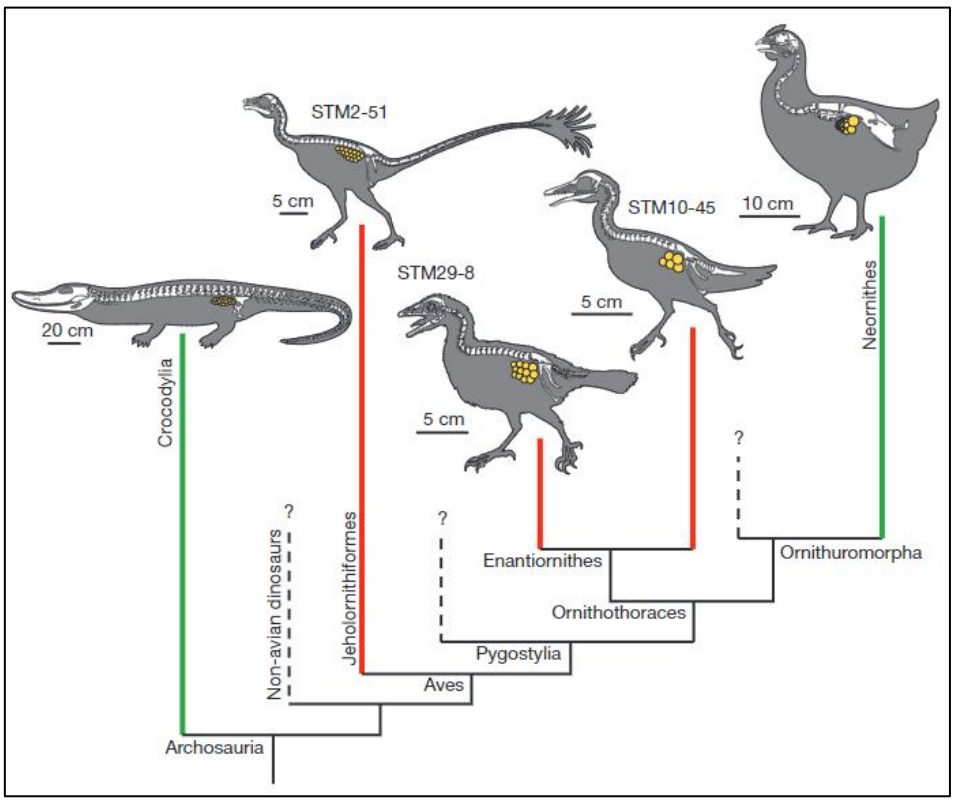


Fig 2.1 Evolutionary lineage of chicken ovaries (Zheng *et al.*, 2013).

In females, substantial expression of the aromatase gene (around day 5–6) leads to estrogen production and selective expression of the estrogen receptor-mRNA in the left gonad leads to the creation of a functional left ovary. Other sex variations may be detected in the expression of inhibin subunit genes in chicken embryo gonads, as well as in circulating levels of inhibin, follicle-stimulating hormone (FSH), and steroids (Bruggeman *et al.,* 2002). At embryonic 4.5 days, the left and right sides of the female chicken embryos initially had the same gonadal morphological appearance. The left gonads of the female chicken embryos were noticeably bigger than the right gonads on the embryonic 7^{th} day as development progressed. As a result, by the embryonic 10^{th} day, the left gonad's dimensions were over 2.5 times larger than the right gonad's (Li *et al.,* 2022).

The left gonad has considerably more germ cells than the right at stage 35. The expression pattern of ERNI (Ens1), a gene expressed in chick embryonic stem cells downregulated upon differentiation (Intarapat and Stern 2013). The variations in left-right ovarian development are determined by the unique expression of oestrogen receptors in the left gonad, which in conjunction with estradiol production, leads to AMH suppression by estradiol action. The lack of estradiol in the right gonad of a female chick embryo prevents estradiol activity, and hence AMH inhibits cortical development, resulting in right gonad and Mullerian duct regression (Bruggeman *et al.,* 2002).

Left-right morphological asymmetry, with the epithelial layer on the left gonad being thicker than that on the right. By HH29, the epithelium around the right gonad has flattened and taken on a more squamous appearance, emphasizing the left-right asymmetry (Carlon and Stahl, 1985). Myostatin (GDF-8) is detected in different organs of chicken embryos, including the testis and ovary, throughout the embryogenesis stages. Its high levels of expression noticed in the testis and ovary (Kubota *et al.,* 2007). FSH gene expression was lowest in the pituitary of the female embryo on E11 and increased on E17. The expression of FSH in the male pituitary gland did not alter across the days investigated. On E11, FSH mRNA expression was greater in the male pituitary gland than in the female pituitary gland. Luteinizing hormone (LH) mRNA expression in the female pituitary increased on D1 compared to E11(Grzegorzewska *et al.,* 2009).

Quantitative RT-PCR was used to identify the androgen receptor (AR) mRNA level in male and female embryonic gonads, and its expression was greater in females than males at all developmental stages studied. The androgen receptor (AR) was found in the nucleus of cells in the left gonad of female embryos (Tanaka *et al.,* 2017). In early chicken embryos, RNA interference (RNAi) was used to knock down DMRT1. In genetically male (ZZ) embryos, reducing DMRT1 protein expression in ovo causes feminization of the embryonic gonads(Smith *et al.,* 2009).

The research was carried out to look at the histological development of the right and left ovaries in an ostrich embryo (kheirabadi *et al.,* 2014). The ovaries grew unequally, resulting in a bigger left ovary with a visible cortex and medulla. The cortex was made up of germinal cells, germinal epithelium, and somatic cells. The medulla of both ovaries included lacunar channels, blood vessels, interstitial cells, and germ cells (kheirabadi *et al.,* 2014). The right ovary germinal epithelium had a thin layer. The left ovary has a cortex and a medulla, whereas the right ovary has only a medulla and no cortex. Stereological findings demonstrate that as development progresses, the overall volume of all ovarian medulla components increases in the left ovary(González-Morán 2011).

2.2 CHICKEN EMBRYO SEXING

An established model system for researching early vertebrate development is the chicken embryo. The ability to execute changes in ovo, prolong incubation and then track the impact on embryonic development is one of this model's key benefits (Clinton *et al.,* 2016). A PCR-based sexing methodology was developed that combines control and sex-specific responses in a single tube experiment. The test is quick, efficient, and tolerating of low-quality DNA to a broad range of DNA concentrations (Clinton *et al.,* 2001).

XHOL repeat sequences were utilised for the identification of the sex of chicken embryos with 415 bp amplification product along with 18S ribosomal gene as a control gene (Clinton *et al.,* 2001) (Shaikat *et al.,* 2018). SWIM and XHOL genes on chromosome W and DMRT gene on chromosomal Z were found, allowing for unequivocal sex discrimination using PCR and qPCR, respectively.

2.3 NEXT GENERATION SEQUENCING AND BIOINFORMATICS

Early attempts to sequence genes included laborious, time-consuming, and meticulous work. This scenario started to improve in the middle of the 1970s when scientist Frederick Sanger created several quicker, more effective ways of sequencing DNA. Sanger's work in this field was so revolutionary that it won him the 1980 Nobel Prize in Chemistry(Attia and Saeed 2016). Targeted (specific genes) or whole-genome sequencing would be used to characterize the individual patient's disease, and potential treatment modalities would be determined based on NGS sequencing data. The future would involve utilizing the speed and scope of new sequencing technologies and data analysis for diagnosis and treatment(Mardis 2011). Examination of how next-generation sequencing technology and bioinformatics tools are applied to the study of the molecular evolution of aquatic species and talk about the present issues and potential future uses of bioinformatics in efforts to conserve aquatic creatures (Tan *et al.*, 2019). A better picture of the animal virome is now available thanks to the recent advancement of unbiased metagenomic next-generation sequencing, which is also throwing fresh light on viral evolution during the COVID 19 pandemic (Harvey and Holmes 2022).

2.3.1 Evolution of sequencing technology

Sequencing speed, read duration, and throughput have all greatly improved over the previous ten years, and the cost per base has dropped significantly as well (Van *et al.*, 2014). DNA sequencing has previously been done using conventional techniques such as Sanger's chain termination method and the Maxam-Gilbert chemical degradation method(Bansal *et al.*, 2018). The first human genome sequence was finally completed by using the Sanger technology (Van *et al.*, 2014). Pyrosequencing, developed by 454 Life Sciences (now Roche), was the first NGS technology to be made publicly available in 2005. The Solexa/Illumina sequencing platform was launched a year after Pyrosequencing and this technology was acquired by Illumina in 2007(Van *et al.*, 2014). The third technology was Sequencing by Oligo Ligation Detection (SOLiD), which was introduced in 2007 by Applied Biosystems (now Life Technologies). Ion Torrent (now Life Technologies) was developed by Jonathan Rothberg in 2010. Compared to other sequencing technologies, Illumina sequencing

produces a significantly bigger number of sequences, with better coverage, and at a lower cost(Yu *et al.*, 2017).

When compared to conventional Sanger sequencing, next-generation sequencing methods enable significantly more rapid and affordable generation, revolutionizing genomics research and other fields (Marguerat and Bähler 2010). Traditional sequencers are unable to read an organism's entire genome in a single run. Instead, in a typical NGS run, thousands or even millions of short overlapping sequences are generated simultaneously(Bansal *et al.*, 2018). NGS techniques are not only applicable to whole genome sequencing, but also to transcriptome sequencing (RNA-Seq), whole exome sequencing (WES), candidate gene sequencing (CGS), genotyping by sequencing (GBS), and chromatin immunoprecipitation sequencing, also known as Chip-Seq(Bansal *et al.*, 2018). Effective algorithms or software that can convert sequence readings and data into useful information are highly needed. Innovative bioinformatics analytic techniques are required, as well as infrastructure to store the abundance of data that is generated day by day (Bansal *et al.*, 2018).

2.3.3 Next generation sequencing – transcriptome sequencing (RNA-Seq)

The technique known as RNA-seq makes use of next-generation sequencing technologies to analyse dynamic transcriptomes in addition to static genomes (Marguerat and Bähler 2010).

For whole-genome transcriptome profiling, RNA-Seq, which is the direct sequencing of transcripts using high-throughput sequencing technologies, has the potential to replace microarrays (Mortazavi, 2008; and Mutz *et al.*, 2013).

RNA preparation, cDNA synthesis, cDNA library generation, cDNA sequencing utilizing a Next Generation Sequencing (NGS) technology, and bioinformatics analysis of the sequence reads produced are all steps in a typical RNA-seq analysis workflow. Assembling reads into transcripts or matching them to reference sequences, labelling putative transcripts, counting the number of reads per transcript, and statistically comparing transcript abundance across samples or treatments are typical steps in bioinformatics analysis (Bullard *et al.*, 2010;).

Table 2.1 Details of the different platforms available for next-generation sequencing (Gullapalli *et al.*, 2012; Fox *et al.*, 2014; Knief, 2014; Mardis, 2013)

S.No	Company	Platform	Approach	Read length (bases)	Through-put (Gb per run)	Run time	Advantages	Disadvantages
1.	**Roche/454 Life Sciences**	GS FLX Titanium XL+	Pyrosequencing, amplification based on emulsion PCR	700–1000	0.7 Gb	0.35–0.42 days	· Long read assembly allows detection of large structural variations · Short run time	· Lower throughput · Homopolymer errors · Signal interference among neighbouring clusters
		GS FLX Titanium XLR70	-do-					
		GS Junior	-do-	250–400				
2.	**Illumina**	HiSeq 2000	Sequencing by synthesis with a reversible terminator, using bridge amplification	35–100	100–600 Gb	2–11 days	· Ultra high throughput · High capacity of multiplexing	· Short read assembly may miss large structural variations. · Signal interference among neighbouring clusters. · Homopolymer errors
		Genome Analyzer IIx	-do-	35–150	10–95 Gb	2–14 days		
		MiSeq	-do-	35–250	540 Mb–8.5 Gb	4–39 days	· Well-proven sequencing technology. · Fully automated workflow · Low cost. · Fast run time	· Low abundance of amplified template

Table 2.1 (Cont.)

	Company	Platform	Approach	Read length (bases)	Through-put (Gb per run)	Run time	Advantages	Disadvantages
3.	**Life Technologies/ Applied Biosystems**	5500xl SOLiD™ system	Sequencing by ligation, amplification based on emulsion PCR	35–75	120 Gb	7–14 days	· Ultra high throughput · Two-base coding (higher accuracy) · High capacity of · multiplexing	· Short read assembly may miss large structural variations. · Long run time · Signal interference among neighbouring clusters · signal degradation.
		SOLiD™ 4 system	-do-	25–50	25–100 Gb	3.5–16 days		
4.	**Life Technologies/ Ion Torrent**	Ion Proton™ sequencer(Proton I chip)	Ion based semiconductor sequencing	100–400	10 Gb	4 h	· Fast run time · Highly scalable (different chips available) · Low cost	· Newest to the market · Homopolymer errors
		Proton II chip	-do-	100–400	30 Gb	4 h		
		Ion PGM™ sequencer (314 chip)	-do-	100–400	0.01 Gb	1 h		
		Ion PGM™ sequencer (316 chip)	-do-	100–400	0.1 Gb	2 h		
		Ion PGM™ sequencer (318 chip)	-do-	100–400	300 Mb–1 1Gb	3h		
5.	**Helicos BioSciences**	HeliScope™ single molecule sequencer	Single-molecule sequencing	25–55	21–35 Gb	8 days		
6.	**Pacific Biosciences**	PacBio RS	Single-molecule sequencing	250 bp–10 kb				· Overall error rate higher

All transcribed areas of the genome can now be mapped with incredible accuracy. Transcriptome data sequencing has many applications, including fusion discovery, variant detection, transcript annotation, quantification of noncoding RNA and quantification of gene expression(Bansal *et al.,* 2018). With the use of RNA-seq, transcriptomes may be sequenced and quantified with the highest precision and dynamic range, regardless of the size of the individual transcripts (Marguerat and Bähler 2010). Using the RNA-seq method, the expression profiles of mRNAs, lncRNAs, and circRNAs in mid-segments of chicken small intestines infected with E. necatrix were examined and found that 1543 mRNAs (707 upregulated and 836 downregulated), 95 lncRNAs (49 upregulated and 46 downregulated), and 13 circRNAs (9 upregulated and 4 downregulated) were altered by Eimeria necatrix infection (Fan *et al.,* 2020). RNA sequencing (RNA-Seq) was performed to undertake gene expression on skeletal muscle samples at various developmental and growth stages (Ding *et al.,* 2021).

2.3.4 Transcriptome Analysis

Generally, transcriptome analysis is done by two methodologies: genome reference and non-reference genome analyses (Liu *et al.,* 2022). By examining the variations in gene types and expression levels at the systemic level, RNA sequencing (RNA-seq) can directly connect changes in gene expression levels with phenotypic alterations (Ding *et al.,* 2021). Historically restricted to small-scale genetic techniques, disciplines like evolutionary biology or molecular ecology are beginning to include genome-wide analysis due to NGS analysis techniques. A typical RNA-seq workflow includes sampling, RNA extraction, library preparation, sequencing and data analysis (Wolf 2013).

De novo assembly is the process of reconstructing transcripts from reads (or read pairs) in the absence of genomic or transcriptome information. Reads can directly match onto the reference when transcript or genome information is easily available (Wolf 2013).

Transcript abundance and isoform identity varies significantly among tissues and undergo dramatic changes not just throughout embryological development but also over the course of an individual's lifetime (e.g., depending on their reproductive status)

and even throughout the course of a day (circadian rhythms). These transcript or mRNA changes can be studied through transcriptome data analysis (Wolf 2013).

2.3.5 Transcriptome Analysis Computational tools

Machine learning algorithms, such as neural networks and support vector machines, as well as recent advances in artificial intelligence, will be crucial in enhancing NGS platforms and software, which will aid researchers and medical professionals in overcoming difficult biological problems (Pereira *et al.,* 2020). The need for statistical approaches or software to evaluate quantitative differences between trials or genes has arisen as a result of the fast adoption of high-throughput sequencing (HTS) technologies for genomic studies (Love *et al.,* 2014).

2.3.5.1 FastQC

FASTQ has become a widely used file format for transmitting sequencing read data that include both a sequence and an associated per base quality score (Cock *et al.,* 2009). Before downstream analysis like mapping or assembly of sequenced raw data (FASTQ) must be pre-processed, because they include adaptor sequences, duplicate quantification, etc. Since the effectiveness of NGS analysis depends on the quality of the raw sequence data, various bioinformatic tools, including the NGS QC toolkit, QC-Chain, and FastQC, can assess the quality of raw data. One of the most popular is FastQC (Pereira *et al.,* 2020). Standard tools like the fastQC can carry out simple tasks like statistics on the quality score(Wolf 2013).

In the study of comparative sequence analyses of the genome and transcriptome variants in the Asian elephant Elephas, Raw FASTQ sequenced reads were checked using the FastQC tool (Reddy *et al.,* 2015). FastQC programmed is advised for quality check analysis (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC to aggregate the reports that are generated (Batut *et al.,* 2021).

2.3.5.2 Trimmomatic

Trimmomatic is a versatile trimmer for Illumina sequences read data. Trimmomatic (Java 1.5 is necessary) has a GPLV3 license. It is available at http://www.usadellab.org/cms/index.php?page1&4trimmomatic (Bolger *et al.,* 2014).

Trimming Illumina data can be done using several tools, such as BTrim, IeeHom, AdapterRemoval, and Trimmonatic (Pereira *et al.,* 2020). Read data trimming and filtering, identifying adapter sequences, and quality filtering are a few of the processes that are included in Trimmomatic. Its performance has been compared with other trimming tools and Trimmomatic was found to give results that were at least competitive with and even better than those produced by other trimming tools (Bolger *et al.,* 2014). Trimmonatic has been used for processing raw data for examination of the expression profiles of mRNAs, lncRNAs, and circRNAs in mid-segments of chicken small intestines infected with E. necatrix (Fan *et al.,* 2020).

2.3.5.3 Hierarchical indexing for spliced alignment of transcripts (HISAT2)

An effective general-purpose tool for mapping sequence readings to genome data is HISAT2 (hierarchical indexing for spliced alignment of transcripts) (Kim *et al.,* 2015). Aligners work by mapping reads of sequenced DNA or RNA to the appropriate locations in a reference genome(Musich and Gosnell 2020). To discover which aligner performed the best on the presented dataset, each aligner was compared against one another. BWA, which had the highest alignment rate, and HISAT2, which had the quickest runtime, were the two best aligners. Since both aligners exhibited comparable transcriptome coverage regardless of alignment rate, HISAT2 was shown to be the superior aligner overall (Musich and Gosnell 2020). (Nie *et al.,* 2022) utilised HISAT2 v2.2.1 during transcriptome studies of Chicken Follicles, the clean reads were mapped to the reference chicken genome (Ensembl *Gallus-gallus*-6.0).

Using HISAT2 v2.1.0, paired-end clean reads were aligned to the reference genome during the transcriptome analysis of important genes and pathways linked to egg production in Nandan-Yao domestic chicken (Sun *et al.,* 2021). StringTie-Ballgown yields the fewest DEGs when the same thresholds for fold change and p-value are used in DE analysis, but HTseq-DESeq2, -edgeR, or -limma typically create more DEGs (Liu *et al.,* 2022) The three most widely used alignment programs are TopHat, HISAT, and STAR, all of which need a reference genome. Currently, HISAT is a commonly used program and is an upgraded version of TopHat (Liu *et al.,* 2022). HISAT2 can save more time and internal storage than Tophat due to the bwt algorithm and FM index (Wen 2017). BWA, which had the highest alignment rate, and HISAT2, which had the quickest runtime, were the two best aligners. Since both aligners

exhibited comparable transcriptome coverage regardless of alignment rate, HISAT2 was shown to be the superior aligner overall (Musich and Gosnell 2020). HISAT2 is used for mapping trimmed clean reads with the *Gallus gallus* reference genome during expression studies of mRNAs, lncRNAs, and circRNAs in mid-segments of chicken small intestines infected with E. necatrix (Fan *et al.,* 2020). For identification of long non-coding RNAs (lncRNAs), messenger RNAs (mRNAs), and microRNAs (miRNAs) during gonadal development in 6th day embryonic chickens, HISAT2 was used to map the clean reads to the chicken reference genome, GRCg6a (GCA 000002315.5) (Zou *et al.,* 2020).

2.3.5.4 featureCounts

featureCounts is a compact read-counting application run in C language. It may be used to count both gDNA-seq and RNA-seq reads provided in SAM/BAM files. featureCounts is available as part of the Subread (http://subread.sourceforge.net) or Rsubread (http://www.bioconductor.org) software packages under the GNU General Public License (Liao *et al.,* 2014). The files of aligned reads in either Sequence Alignment/Map (SAM) or Binary Alignment/Map (BAM) format are provided as the input to featureCounts along with a genome annotation file in either general feature formats (GFF) or Gene transfer format (GTF) file format (Liao *et al.,* 2014).

featureCounts assigns read precisely by comparing the mapping location of each base in the read or fragment to the genomic region. Any gaps (insertions, deletions, exon-exon junctions, or fusions) discovered during the reading are taken into consideration (Liao *et al.,* 2014). The featureCounts was discovered to be faster efficient than other counting tools. Due to its extremely quick search method and highly effective implementation using the C programming language, featureCounts has a great computational efficiency (Liao *et al.,* 2014).

2.3.5.5 DESeq2

Differential gene quantification between samples can be performed by the statistical tool DESeq2, which is based on the R-language. It requires read counts, which we can rely on to set up a matrix and use contraction equation estimation and negative binomial distribution to gain insights into the variations in gene expression

(Wen 2017). Differentially expressed genes (DEGs) were identified using DESeq2 v1.32.0 during transcriptome studies of chicken follicles (Nie *et al.*, 2022).

The DESeq2 software was used to undertake a differential expression study of transcripts from theca cells from two different types of chicken follicles (Shen *et al.*, 2020) The differentially expressed genes were found using the DESeq2 R-package (1.18.0) during transcriptome analysis of important genes and pathways linked to egg production in Nandan-Yao domestic chicken (Sun *et al.*, 2021). For the identification of long non-coding RNAs (lncRNAs), messenger RNAs (mRNAs), and microRNAs (miRNAs) during gonadal development in sixth-day embryonic chickens, DESeq2 was used to identify DEmRNAs and DElncRNAs that differed between samples (Zou *et al.*, 2020). DESeq2 combines methodological innovations with several additional features to allow for a more quantitative examination of comparative RNA-seq data utilizing shrinkage estimators for dispersion and fold change (Love *et al.*, 2014).

2.3.5.6 DESeq2 Shiny

DESeq2 Shiny is a part of Nucleic Acid Sequence Analysis Resource or NASQAR, which is a web-based high-throughput data analysis and visualization (Yousif *et al.*, 2020).

2.3.5.7 gProfiler

g:Profiler is a publicly accessible web server (http://biit.cs.ut.ee/gprofiler/) for characterizing gene lists obtained from the mining of high-throughput genomic data. Gene Ontology (GO), pathway, or transcription factor binding site enrichments can be captured using g:Profiler's user-friendly web interface and strong visualisation up to the individual gene level. The Ensembl database is one of the sources used to update the underlying data for g:Profiler, which covers 31 distinct species. (Reimand *et al.*, 2007).

2.3.5.8 ShinyGo

ShinyGO (http://bioinformatics.sdstate.edu/go/) was developed based on a large annotation database obtained from Ensembl and STRING-db for 59 plants, 256 animals, 115 archeal, and 1678 bacterial species. ShinyGO's innovative features are the graphical representation of enrichment results and gene properties, as well as the

application providing interface access to KEGG and STRING for the retrieval of route diagrams and protein-protein interaction networks (Ge *et al.,* 2020).

2.3.5.9 Galaxy Server

Galaxy (https://galaxyproject.org/) is an open-source, web-based platform for computational biomedical research that is accessible, reproducible, and transparent. It enables users without programming knowledge to quickly input parameters and run both smaller workflows and individual tools. Additionally, it records run details so that any user can reproduce and comprehend a thorough computational study. The ever-increasing computational requirements of biomedical data analysis have been the impetus behind ongoing advancements in Galaxy's heterogeneous computing infrastructure utilization (Jalili *et al.,* 2020). Galaxy server can also be installed on local servers or computers using the instructions provided at https://crs4.github.io/Galaxy4Developers/lectures/05.get_galaxy_up_and_running/.

2.4 SEQUENCE READ ARCHIVE (SRA) DATA BANKS

The principal next-generation sequence data archive, the Sequence Read Archive (SRA), has seen accelerated growth in the amount of data provided due to the combination of substantially lower cost and faster sequencing techniques (Leinonen *et al.,* 2010).

The International Nucleotide Sequence Database Collaboration (INSDC) maintains the SRA, which was created as a public repository for next-generation sequence data. The National Center for Biotechnology Information (NCBI), the European Bioinformatics Institute (EBI), and the DNA Data Bank of Japan (DDBJ) are some of the collaborators of INSDC (Leinonen *et al.,* 2010). The SRA can be accessed through NCBI, EBI, and DDBJ at https://www.ncbi.nlm.nih.gov/sra, https://www.ebi.ac.uk/ena/browser/home, and https://www.ddbj.nig.ac.jp/dra/index-e.html respectively.

2.5 STUDIES ON CHICKEN OVARY

During chicken embryonic development (E10–E18), increased expression of FSHR and CYP19A1 was found in the left ovary as compared to the right ovary,

favouring left ovarian development and functionality(Shaikat *et al.,* 2018). Reverse transcriptase polymerase chain reaction (RT-PCR) was used to determine the amount of P450 arom mRNA in the left and right presumptive ovary and testis of growing chicken embryos and found that during initial morphological gonadal differentiation asymmetric function of the P450 arom gene expression was seen (Villalpando *et al.,* 2000).

Identification of long non-coding RNAs (lncRNAs), messenger RNAs (mRNAs), and microRNAs (miRNAs) during gonadal development in chick embryos on the sixth-day was performed and found that for the sex-biased genes, a total of 8 DEmiRNAs, 183 DElncRNAs, and 123 DEmRNAs were discovered; for the side-biased genes, a total of 7 DEmiRNAs, 189 DElncRNAs, and 183 DEmRNAs were discovered (Zou *et al.,* 2020). RNA sequencing was done to profile circRNAs and mRNAs in theca cells from three different types of follicles: small yellow follicles (SYF), the smallest hierarchical follicles (F6), and the largest hierarchical follicles (F1) and found differential expression of 14,502 circRNAs in the theca cells (Shen *et al.,* 2020).

For understanding right ovary degeneration in chicken, high-throughput RNA-sequencing technologies were used and identified 539 and 1046 genes that showed significantly different expression between embryonic 6^{th} day and 10^{th} day respectively (Yu *et al.,* 2017). In chickens, western blot analysis was used to determine the levels of androgen receptor (AR) in the left and right ovaries and found that more content of androgen receptor (AR) is present in the left ovary (González-Morán *et al.,* 2013).

Certain reproductive anomalies, including atresia and a double ovary-oviduct, were discovered during a QTL study involving an intercross between White Plymouth Rock chicks and Red Jungle Fowl(Sutherland *et al.,* 2017). The sequence of histological and stereological changes in various parts of the left developing and right regressing ovaries of Chicken was examined. It was discovered that the left ovary has a cortex and a medulla, whereas the right ovary only has a medulla and no cortex (González-Morán *et al.,* 2013). Ostrich's embryonic right and left ovaries' histological development was investigated, and it was discovered that the right ovary lacked a cortex and had a thin layer of germinal epithelium (Nabipour *et al.,* 2015). Studies on the relative increase in the mass and length of the chicken embryo and its reproductive organs were conducted,

and it was observed that during the 12th and 13th days of embryonic development, chicken embryos gained the maximum weight and ovarian mass (Khokhlov *et al.,* 2020).

The asymmetric expression pattern of BMP7 during chicken gonadogenesis in the left-right ovaries is sex-specific during genital ridge formation. In experimentally induced female-to-male reversal with the aromatase inhibitor, ovary-specific BMP7 expression was decreased (Hoshino *et al.,* 2005). In the earlier and later stages, the left ovary had higher levels of BMP3 expression than the right. Early ovarian stages showed that CGNRH-R expression was more prominent in the left ovary, although NR5A2 and GREM1 expression was more in the right ovary. Later ovarian stages showed three genes SOX9, IGF1, and GAL were increased on the left whereas CYP11A1, FST, and PAX2 were expressed more on the right. However, some genes like NR5A2, NROB1, CYP11A1, EMB, FST, GREM1, and PAX2, are elevated in the right ovary, which might serve as ovarian regression indicators (Carré *et al.,* 2011). Transcripts such as TSPAN1, TSPAN8, GIIEp, galectin-2, gastrotropin, and SOX9 have expression patterns that support the growth of the left ovary (Scheider *et al.,* 2014).

2.6 HOUSEKEEPING GENES

Housekeeping genes (HKGs) are employed for normalization in expression research because they provide stability and unregulated expression in the sample type under study (Romanowski *et al.,* 2007). A reference gene is necessary to address fundamental sample differences, such as variations in cellular input, RNA quality, reverse transcription efficiency, and batch-to-batch variance in reagents. The typically utilized housekeeping genes include glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β-actin, β2-microglobulin, cyclooxygenase 1, hypoxanthine phosphoribosyl transferase 1, glucose-6-phosphate dehydrogenase, tubulin, transferrin receptor, and 18S ribosomal RNA (Lee *et al.,* 2002).

The β-actin gene, which produces a structural protein of the cytoskeleton, is the most frequently utilized gene for normalizing gene expression data(Pohjanvirta *et al.,* 2006). β-Actin demonstrated good stability in both fresh and *in-vitro* produced goat preantral follicles (Frota *et al.,* 2011). In studies using lymphoid cells from cattle, β-actin was steady when measured using the NormFinder tool (Anstaett *et al.,* 2010).

2.7 REAL-TIME PCR

Real-time PCR is increasingly being used for genetic analysis and RNA quantification (Ponchel *et al.,* 2003). The ability to detect and precisely measure the target genes, even at low expression levels, makes it the most useful technique for analysing gene expression modulations (Tajadini *et al.,* 2014). Its strength lies in its capacity to use fluorescence to measure the amount of PCR product (amplicon) at each cycle of the PCR (Ponchel *et al.,* 2003). Two methods are frequently used in quantitative gene expression analysis: the SYBR Green method, which uses double-stranded deoxyribonucleic acid (dsDNA) to bind the fluorescent dye, and the TaqMan method, which uses a dual-labelled oligonucleotide and the exonuclease activity of the Taq polymerase enzyme to measure the amount of amplified product in real-time (Tajadini *et al.,* 2014).

The four phases of RT-PCR are the linear ground phase, early exponential phase, log-linear phase (also known as exponential phase), and plateau phase (Tichopad *et al.,* 2003). Under ideal circumstances, RT-PCR takes use of the fact that the amount of PCR products in the exponential phase is proportional to the amount of initial template (Gibson *et al.,* 1996; and Heid *et al.,* 1996). The fluorescence emission at each cycle has not increased above the background level during the linear ground phase. At the exponential phase, the amount of fluorescence reaches a threshold where it can be identified as being noticeably stronger than the background fluorescence signal. According to the documentation for ABI PRISM® (Applied Biosystems, Foster City, CA, USA) and LightCycler (Roche Applied Science, Indianapolis, IN, USA), the cycle in which this detection takes place is referred to as the Threshold cycle (CT) or crossing point (CP) (Heid *et al.,* 1996; and von Ahsen *et al.,* 1999). When calculating experimental results, the CT value serves as a representation of the copy number of the original template (Heid *et al.,* 1996).

2.7.1 Quantification method

Absolute and relative quantitation are two quantification techniques utilized to examine qPCR data. In order to create a standard curve for absolute quantification, standards with known concentrations are serially diluted. A standard curve between CT and initial amounts of total RNA or cDNA results in a linear relationship, allowing the

concentration of unknowns to be determined based on their CT values (Heid *et al.,* 1996). A reference sample or an external standard is used to assess changes in sample gene expression in relative quantitation (calibrator). $2^{-\Delta\Delta CT}$ is a practical technique for assessing the relative expression of genes in qPCR (Livak and Schmittgen 2001).

Fold change = $2^{-\Delta\Delta CT}$

where, ΔΔ CT = Δ CT (sample) – Δ CT (reference) (Rao *et al.,* 2013).

2.7.2 Advantages of Real-time PCR

RT-PCR can generate quantitative data with a precise dynamic range of log 7 to log 8 orders of magnitude (Morrison *et al.,* 1998). In comparison to RNase protection assay (Wang *et al.,* 1999) and dot blot hybridization, RT-PCR assays are 10,000 to 100,000 fold and 1,000 folds more sensitive, respectively (Malinen *et al.,* 2003). Even a single duplicate of a certain transcript can be found using RT-PCR (Palmer *et al.,* 2003). Due to its closed system and lack of post-PCR processing, it prevents the possibility of PCR product carryover and enables experiments with a much greater throughput (Heid *et al.,* 1996). When compared to other techniques for gene expression analysis, it can also discriminate between messenger RNAs (mRNAs) with identical sequences and also uses a lot less RNA template (Wong *et al.,* 2005).

MATERIALS AND METHODS

CHAPTER - III

MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 Experimental animals

Kadaknath chicken maintained at ICAR- Directorate of Poultry Research, Rajendranagar, Hyderabad was used for the present study. The female birds were inseminated with Kadaknath semen. The eggs laid were collected and kept for incubation at a temperature of 99.5^{o} F and relative humidity of 55-60%. On the 12th day of incubation, candling of eggs was done and fertile eggs were collected for experimental study.

3.1.2 Collection of samples

3.1.2.1 Tissue

Twenty fertile eggs of Kadaknath birds on the 12th day of incubation were randomly chosen for the collection of tissue samples. Left and right ovaries (Fig. 3.4 and Fig. 3.5) and liver were collected from female embryos to study the expression profile of up and down-regulated genes involved in regression of the right ovary. The embryos were sacrificed and the tissues were collected aseptically and immediately submerged in the RNA later solution for storage. The samples were brought to the laboratory and stored at -80ºC till the isolation of RNA from the ovary and DNA from the liver.

3.1.2 Glass and Plasticware

For RNA extraction, the work surface of the laboratory and laminar flow cabinet were wiped with 70% ethanol. On the day of isolation, the UV lamp of laminar flow was switched on for 20 min before the processing of tissue. The mortar and pestle used for isolation were cleaned and rinsed with diethylpyrocarbonate (DEPC) treated water (Appendix I). It was mixed thoroughly on a magnetic stirrer at 37°C and autoclaved to remove all traces of DEPC, wrapped in aluminium foil and then autoclaved again.

Fig. 3.1: (a) Kadaknath flock (b) Kadaknath Male chicken (c) Kadaknath female chicken

Fig. 3.2: Infertile egg

Fig. 3.3: Fertile egg

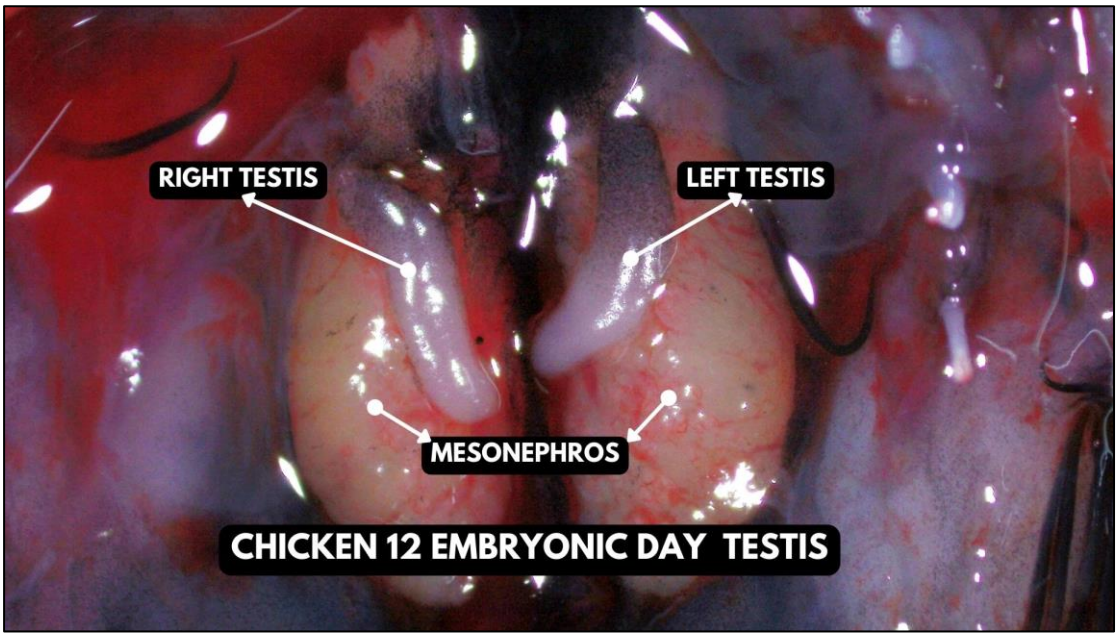


Fig. 3.4: Embryonic 12th day Chicken Testis

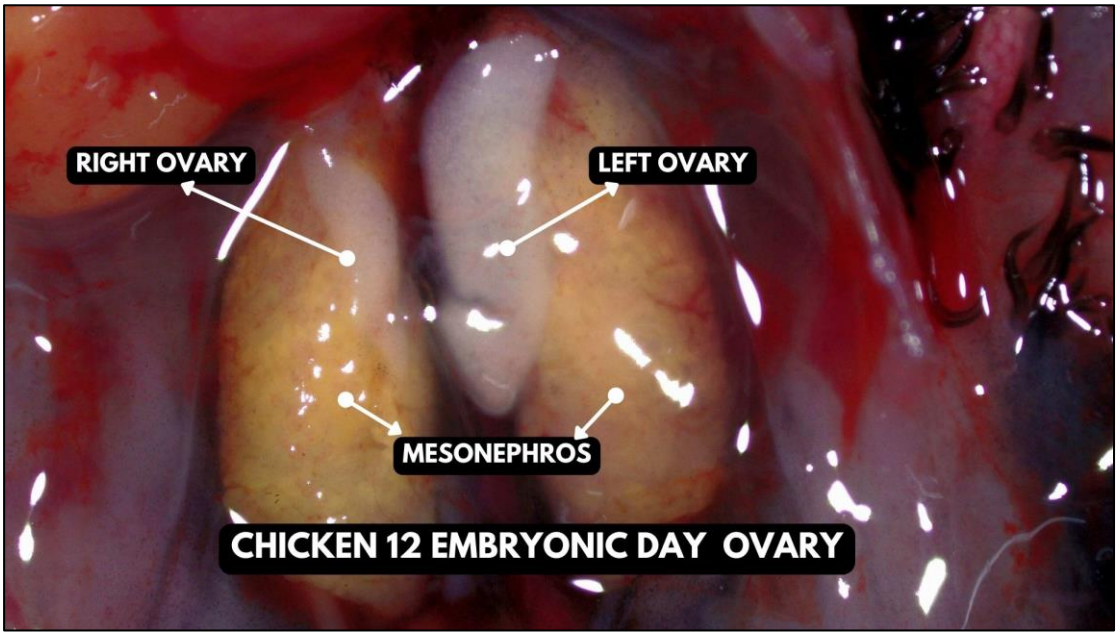


Fig. 3.5: Embryonic 12th day Chicken Ovaries

Scissors and forceps were autoclaved before use. The plastic wares used were treated with 0.1% DEPC to make them DNase-RNase free, autoclaved to neutralize the activity of DEPC and dried finally. RNaseZAP (Sigma) was used to clean the gloves and work surface to ensure RNase free environment for RNA isolation. For PCR, all the glassware was washed and sterilized by following standard lab procedures and all plasticware was autoclaved (121ºC for 20 min at 15 psi pressure).

3.1.2 RNA-Seq Dataset

Data sets needed for the identification of differently expressed genes between the left and right ovary were found in the NCBI SRA database (Table 3.1) . The libraries which were found in the NCBI SRA database were sequenced by Annoroad (Beijing, China) on an Illumina NextSeq 500 platform (Illumina, Santiago, CA, USA), and 50 bp single-read reads were generated (Wan *et al.*, 2017).

Table 3.1 Details of the RNA-Seq Datasets used in the study

S.No	Sample	Assigned Name	NCBI SRA No
1.	Embryonic 12th Day Right Ovary Sample 1	E12FR_1	SRR4029463
2.	Embryonic 12th Day Right Ovary Sample 2	E12FR_2	SRR4029464
3.	Embryonic 12th Day Left Ovary Sample 1	E12FL_1	SRR4029461
4.	Embryonic 12th Day Left Ovary Sample 2	E12FL_2	SRR4029462

3.1.3 Computational resources

The Galaxy US Server instance which is a cloud platform running on different computing clusters (Jalili *et al.*, 2020), which provides the necessary tools along with storage and compute infrastructure, was used for the DEGs analysis.

3.1.4 Softwares

The study utilised different software for the analysis of transcriptome data to the identification of differentially expressed genes. Quality check of SRA was performed by FastQC (Version 0.11.9) tool, Trimmomatic (Version 0.38) used for quality filtering and read trimming, HISAT2 (Version 2.2.1) for mapping, featureCounts (Version 2.0.1) for gene quantification, DESeq2 (Version 2.11.40.7) for differential gene expression, Volcano plot (Galaxy Version 0.0.5) using ggplot2, Microsoft Excel for statistical analysis.

3.2 TRANSCRIPTOME ANALYSIS

3.2.1 Upload FASTQ datasets to Galaxy

Raw data from the NCBI SRA in FASTQ format is typically the starting point of an RNA-Seq investigation. A new history for this RNA-Seq exercise is created and SRA FASTQ files mentioned in Table 3.1 are uploaded into the galaxy server. Uploading data sets was performed based on the fasterq-dump utility of the SRA Toolkit. This tool extracts data (in fastq format) from the Short Read Archive (SRA) at the National Center for Biotechnology Information (NCBI).

3.2.2 Quality Control and Trimming

Errors are encountered during sequencing, such as faulty nucleotide counts due to sequencing platform technological limitations. Sequencing mistakes may influence the analysis and cause the results to be incorrectly interpreted. Sequencing adapters need to be removed before downstream processing. Therefore, the initial step in every analysis should be sequence quality control. All the fastq datasets in the study are given as input to the FastQC tool and any errors in the report were rectified using Trimmomatic. The Average quality required of the output fastq files was set as 30 in Trimmomatic operation.

3.2.3 Mapping reads to the reference genome

The clean reads obtained after quality check and trimming are further processed for mapping with the reference genome. *Gallus gallus* (bGalGal1.mat.broiler.GRCg7b) reference genome in fna.gz format is used as input to HISAT2 along with clean reads.

The *Gallus gallus* reference genome which is available at the NCBI FTP server (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/016/699/485/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b_genomic.fna.gz) is utilised for mapping. HISAT2 generated mapped output files in bam format.

3.2.4 Quantification of annotated genes

Quantifying the reads per gene, or more precisely, the reads that map to each exon is an initial step in comparing the expression of individual genes under various situations. featureCounts is a quick and effective tool for this operation. *Gallus gallus* (bGalGal1.mat.broiler.GRCg7b) genome annotation in gtf.gz format along with HISAT2 output files in bam format are used as input to featureCounts. The *Gallus gallus* genome annotation which is available at the NCBI FTP server (https://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_other/Gallus_gallus/latest_assembly_versions/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b_genomic.gtf.gz) is utilised for quantification. featureCounts generated gene quantification data in tabular format. The output of featureCounts contains a table that lists the number of reads that have been mapped to each gene in the given *Gallus gallus* genome annotation.

3.2.5 Differential gene expression analysis between samples

DESeq2 is a program that employs negative binomial generalized linear models to perform differential analysis on output files of featureCounts. DESeq2 normalizes for sequencing depth and library composition after combining read count data from many samples into a large table. Since we are comparing counts for the same gene across sample groups, gene length normalization is not necessary. The tabular count files from the featureCounts are given input to DESeq2. The input parameters for the DESeq2 factor name were given as right_left_ovary. Factor level 1 or treatment group contains featureCounts output count files of the right ovary where as Factor level 2 or control group contains featureCounts output count files of the left ovary, which resulted in an output of DESeq2 result file containing differentially expressed genes between left and right ovary.

3.2.6 Filtering significant up and down-regulated genes

DESeq2 output file contains both significant and non-significant differentially expressed genes. Differentially expressed genes with a p-value below 0.05 are filtered out in Microsoft excel and they are considered to be significantly differentially expressed genes. Up-regulated genes are filtered out from significant differentially expressed genes having fold change (log2FC) above +2. Down-regulated genes are filtered out from significant differentially expressed genes having fold change (log2FC) below -2.

3.2.7 Functional enrichment analysis

To determine which biological activities may be involved by the differentially expressed genes, enriched transcripts of genes that belong to more general or specialized categories need to be known. In genome-wide expression studies, gene ontology (GO) analysis is frequently used to simplify and emphasize biological processes. Gene ontology (GO) analysis is carried out using ShinyGO which uses a number of R or Bioconductor programs and a huge annotation-pathway database that was gathered from various sources to understand pathways involved by differentially expressed genes. g:Profile, a simple user-friendly application, was also used for further analysis and visualisation of Gene Ontology (GO) and pathways.

3.3 SEXING OF CHICKEN EMBRYOS

Identification of sex of embryos is needed to differentiate collected gonads into ovary and testis, so that ovary samples are further processed to carry out expression studies and testis samples are discarded. Chicken embryo sexing is done by identification of the SWIM gene on chromosome W using PCR. DNA is extracted from liver samples of embryos.

3.3.1 Isolation of Genomic DNA

Genomic DNA was isolated from the chicken embryo liver by phenol-chloroform extraction method by standard protocol (Sambrook *et al.*, 1989) with slight modifications. All the chemicals and reagents used for the DNA isolation are enlisted

in Appendix I. The steps followed in the isolation of DNA from chicken embryo liver samples are detailed below.

Step 1: Tissue sample is taken, add 500 ml PBS and homogenize in mortar & pestle.

Step 2: Transfer homogenised mixture into 2ml tube and spin at 2000 RPM for 5 min.

Step 3: Decant supernatant and take 100 ml of suspended tissue.

Step 4: Add 900 ml of tissue lysis buffer, 15 ml of proteinase K (25 µg/µl) and 100 ml of SDS and incubate overnight at 37°C in a water bath.

Step 5: After overnight incubation add 900 ml of Tris-saturated Phenol (pH>8) and spin at 10,000 RPM for 10 min.

Step 6: Transfer the upper aqueous phase containing DNA into another 2ml tube and add 900 ml of Phenol: Chloroform: Isoamyl alcohol (25:24:1).

Step 7: Spin the contents at 10,000 RPM for 10 min.

Step 8: Transfer the supernatant into another tube and add 900 ml of Chloroform: Isoamyl alcohol (24:1).

Step 9: Spin at 10,000 RPM for 10 min.

Step 10: Transfer the supernatant into another tube and add an equal volume of isopropyl alcohol and mixed by gentle inversion and kept at room temperature for 5 min for precipitation of DNA.

Step 11: Spin at 5000 RPM for five minutes and let the pellet settle down and the supernatant was discarded by inversion.

Step 12: The DNA pellet was washed in 70% ethanol twice. Centrifuged at 10,000 RPM for 10 min. The supernatant was discarded by inversion.

Step 13: Finally, the DNA pellet was air dried by inverting the tube onto blotting paper and then the pellet was dissolved in 200 µl of nuclease-free water (NFW). Store the extracted DNA at 4°C.

3.3.1 Evaluation of Purity, Quality and Concentration

3.3.1.1 Purity of DNA

The purity of the genomic DNA was assessed by checking the optical density (OD) at 260 nm and 280 nm which indicates the amount of DNA and amount of protein respectively in a given sample. The ratio of optical densities at 260 nm and 280 nm was used as a criterion for purity. The samples having an OD ratio (260 nm/280 nm) of 1.7 to 1.9 were used for the experiment.

3.3.1.2 Concentration of DNA

The concentration of genomic DNA was estimated by taking the OD value at 260 nm. The genomic DNA concentration was measured by using the formula given as

DNA concentration (µg/ml) = OD x Dilution factor x 50

3.3.1.3 Quality of DNA

Genomic DNA quality was checked to ensure intact DNA without any shearing. Horizontal submarine agarose gel electrophoresis was performed to check the quality of DNA. Agarose of 0.8% w/v was dissolved in 1x TBE buffer by heating. The agarose solution was cooled and poured into the gel casting tray with a comb after adding ethidium bromide (0.5 µg/ml). After solidification of gel, the comb is removed and submerged in a gel tank having 1x TBE buffer. The isolated DNA samples were mixed with $1/6^{th}$ volume of 6x gel loading buffer and loaded into the wells using a micropipette. The electrophoresis was carried out at 70 volts for about 1h. Then the gel was visualized under UV light and photographed using a gel documentation system. After checking the quality, the isolated DNA was diluted to 50-100 ng/µl with nuclease-free water (NFW) for further analysis.

Note: Agarose Gel Electrophoresis reagents are presented in Appendix II

3.3.2 Polymerase Chain Reaction (PCR)

The amplification of the SWIM gene with specific primers (Table 3.2) (Bioserve) was carried out in 0.2 ml capacity PCR tubes, using a thermal cycler (Himedia). A master mix for PCR amplification was prepared as presented in Table 3.3.

Table 3.2 Details of primers used for embryonic chicken sexing

S.No	Targeted gene	Primer sequence	Amplified product	Authors
1	SWIM	F: 5' GAGATCACGAACTCAACCAG 3'	212 bp	(He *et al.*, 2019)
		R: 5' CCAGACCTAATACGGTTTTACAG 3'		
F: Forward, **R:** Reversed				

Table 3.3 Composition of PCR reaction mix for amplification

S.No	Components	Volume	Final concentration
1	10x Taq buffer	1.25 µl	1 X
2	dNTPs (10Mm)	0.5 µl	0.4 mM
3	Primer-Forward (50 pm)	1 µl	4 pm
4	Primer-Reverse (50 pm)	1 µl	4 pm
5	$MgCl_2$(25 mM)	0.5 µl	1 mM
6	Taq Polymerase (1 unit/µl)	0.5 µl	0.5 unit
7	Autoclaved Mille Q water	6.75 µl	

An aliquot of 11.5 µl of master mix per sample was drawn into thin-walled PCR tubes and 1 µl (50-100 ng) of template DNA was added for making 12.5 µl. The PCR tubes were marked for identification, then spun briefly for proper mixing and mounted in the PCR machine.

3.3.2.1 PCR reaction conditions

The PCR protocol with specific annealing temperatures for SWIM primers is detailed in Table 3.4. The annealing temperature is 55°C. The PCR tubes were kept in a thermal cycler and the programme was executed. PCR amplification programme took about 2h. At the end of the PCR, the tubes were taken out and stored at -20°C until

further use. The PCR products were confirmed by agarose gel electrophoresis and visualized in the gel documentation system.

Table 3.4 PCR reaction conditions for SWIM primers

Step	Process	Temperature (°C)	Time
1	Initial denaturation	94	5 min
2	Cyclic denaturation	94	30 sec
3	Primer annealing	55	30 sec
4	Cyclic extension	72	30 sec
5	Steps 2 to 4 were repeated for 35 cycles		
6	Final extension	72	5 min
7	Hold	4	Forever

3.3.2.1 Checking of the amplified product

Horizontal submarine agarose gel electrophoresis (2% w/v) was carried out to check the amplified product. A 5 µl PCR product mixed with 1 µl of 6x gel loading dye was loaded along with a 100 bp ladder as a marker in a separate lane. The electrophoresis was done at 80V for 2h. The amplified product in the gel was observed under a UV transilluminator and documented by a gel documentation system.

3.4 EXPRESSION PROFILING OF DIFFERENTIALLY EXPRESSED GENES

3.4.1 Total RNA isolation using TRIzol reagent

Total RNA was isolated from the samples (ovaries of 12^{th} day chicken embryos) using the TRIzol method by following the manufacturer's protocol.

1. The samples were primarily washed with DEPC treated PBS for removal of excess RNA-later solution and homogenized with micro-pestle, treated with TRIzol reagent (1 ml/50 mg of tissue) and incubated at room temperature for 5 min.

2. About 200 µl of chloroform was added followed by vigorous shaking for 15 seconds and incubated for 2-3 min at room temperature and centrifuged at 12,000 rpm for 15 min at 4°C.
3. Then the upper aqueous phase was collected into a clean eppendorf tube.
4. 500 µl of isopropanol was added to the aqueous phase for subsequent total RNA precipitation. The solutions were gently mixed and incubated for 10 min at room temperature and centrifuged at 12,000 rpm for 15 min at 4°C to obtain the RNA pellet.
5. The resultant pellets were washed in 1 ml 75% DEPC treated ethanol, centrifuged at 7,500 rpm for 5 min at 4°C and dried for 5-10 min followed by re-suspension in nuclease-free DEPC treated water. The final RNA product of each sample was stored at -80°C for their ultimate cDNA synthesis.

3.4.2 RNA quality, quantity and purity

The quality, quantity and purity of isolated RNA samples were measured by Nanodrop™ 1000 spectrophotometer (Thermo Scientific, USA). The ratio between the OD measurements at 260 nm and 280 nm was used to estimate the purity of the RNA stock. A good quality sample with an OD value ranging from 1.9 to 2.2 was used for cDNA synthesis.

3.4.3 cDNA synthesis

cDNA was synthesized from total RNA samples by using a high-capacity Puregene First Strand cDNA Synthesis Kit (Genetix Biotech) which is a complete system for efficient synthesis of first strand cDNA from mRNA templates. The reagents needed to prepare the master mix were thawed and the master mixture was prepared in ice following the manufacturer's protocol. The kit uses Reverse Transcriptase (RT), which has lower RNase H activity compared to AMV reverse transcriptase. Protocol for First Strand cDNA Synthesis is as follows:

Step 1: Add the reagents in the indicated order into a sterile, nuclease-free tube on ice as mentioned in Table 3.5.

Step 2: The RNA template is GC-rich or contains secondary structures, incubate at 65°C for 5 min on Thermocycler and place back on ice.

Step 3: Add the components in Table 3. First add 4 µl 5X reaction buffer, then add 1 µl RNase Inhibitor (20U/µl), then add 2 µl 10 mM dNTP Mix and finally add 1 µl M-MuLV RT (200U/µl).

Step 4: Mix gently, centrifuge briefly and Reverse transcription was done in a thermocycler (Himedia) using the thermal profile given in Table 3.7. The resulting cDNA was then stored at -20°C till use for expression studies.

Table 3.5 Phase I cDNA synthesis reagents added in the indicated order

Reagent Type	Reagent	Volume
Template RNA	Extracted RNA samples	0.1 ng – 5 µg
Primer	oligo$(dT)_{18}$	5 µl
Nuclease-free water		Up to 12 µl
	Final Volume	**12 µl**

Table 3.6: Phase II cDNA synthesis reagents added after primary treatment

S.No	Reagent	Volume
1.	5X reaction buffer	4 µl
2.	RNase Inhibitor (20U/µl)	1 µl
3.	10 mM dNTP Mix	2 µl
4.	M-MuLV RT (200U/µl)	1 µl
Total Volume		**20 µl**

Table 3.7: Thermal profile of cDNA synthesis

S.No	Step	Temperature (°C)	Time
1	Phase I	65	5 min
2	Phase II	42	60 min
3		70	5 min

3.4.4 Optimisation of cDNA with housekeeping gene (β actin and GAPDH genes)

The primers for amplifying chicken β actin and GAPDH genes were designed using Primer 3 software based on available sequences in the GenBank database (NCBI NM_205518 and NM_204305.2). The primers used for amplifying chicken β-actin GAPDH genes are given in Table 3.8. Amplification of these genes from the cDNA samples was performed in a PCR machine for confirmation of cDNA.

Table 3.8: Details of Primers for expression profiling of β actin and GAPDH gene

Gene	Primer	Sequence (5′-3′)	Product size
β actin gene	β actin F	F: TCCCTGGAGAAGAGCTATGAA	109 bp
	β actin R	R: CAGGACTCCATACCCAAGAAA	
GAPDH gene	GAPDH F	F: CTGCCGTCCTCTCTGGC	119 bp
	GAPDH R	R: GACAGTGCCTTGAAGTGT	

3.4.5 Expression profiling

The expression of mRNA of the genes was quantified using Real-Time PCR (Qiagen Rotor-Gene™ Q). The details of primers for validation of selected differentially expressed genes are given in Table 3.9. Chicken β actin and GAPDH were used as a housekeeping gene (endogenous control) for the analysis of data. It is ubiquitously expressed irrespective of a varied environmental condition and/or tissue.

Different component combinations were used to standardize the concentration of each component. The cDNA prepared from left and right ovaries were used as a template. GeneSure™ SYBR Green qPCR Master Mix (Genetix Biotech) was used for the quantification of genes.

Table 3.9: Details of primers for validation of selected differentially expressed genes

S.No	Gene	Primers	Sequence (5′-3′)	Product Size	Reference
1.	cDAZL	cDAZL F	F: TCCCAGAGCCCACACAGATG	160 bp	(Rengaraj *et al.*, 2010)
		cDAZL R	R: AAGTGATGCGCCCTCCTCTC		
2.	cFETUB	cFETUB F	F: CTGACCGAACAGAGGGCTAC	278 bp	
		cFETUB R	R: ATATCTGGGTGGAACTGGCTG		
3.	cDDX4	cDDX4 F	F: GTAGCATCAAGAGGCCTGGA	90 bp	(Aduma *et al.*, 2019)
		cDDX4 R	R: ACGACCAGTTCGTCCAATTC		
4.	cLEAP2	cLEAP2 F	F: CTCAGCCAGGTGTACTGTGCTT	66 bp	(Su *et al.*, 2017)
		cLEAP2 R	R: CGTCATCCGCTTCAGTCTCA		

All the samples or reactions in Real-time PCR were kept in duplicates. The reaction mixture of Real-time PCR is shown in Table 3.10. Thermal cycling conditions used for the qPCR were tabulated in Tables 3.11. The melt curve analysis was carried out to test the specificity of the amplification.

Relative Quantitation (RQ) using the Comparative C_T method (Livak and Schmittgen 2001). The relative quantification was determined by following the formula:

Relative quantification (RQ) = $2^{-\Delta\Delta CT}$

Table 3.10: Reaction mixture of Real-time PCR

S.No	Components used	Volume
1	cDNA template	1 µl
2	SYBR Green qPCR Master Mix	7.5 µl
3	Primer: forward	1.8 µl
4	Primer: reverse	1.8 µl
5	Water	3 µl

Table 3.11: Thermal cycling conditions of Real-Time PCR

Gene	Hold	Cycling			Melt
		Step 1	Step 2	Step 3	
β actin, GAPDH, DAZL, DDX4, and LEAP2	95ºC for 10 min	95ºC for 15 sec	60ºC for 30 sec	72ºC for 30 sec	Ramp from 72°C to 95°C Hold for 90s on the 1st step Hold for 5s on next steps
FETUB	95ºC for 10 min	95ºC for 15 sec	60ºC for 30 sec	72ºC for 30 sec	

RESULTS

CHAPTER IV

RESULTS

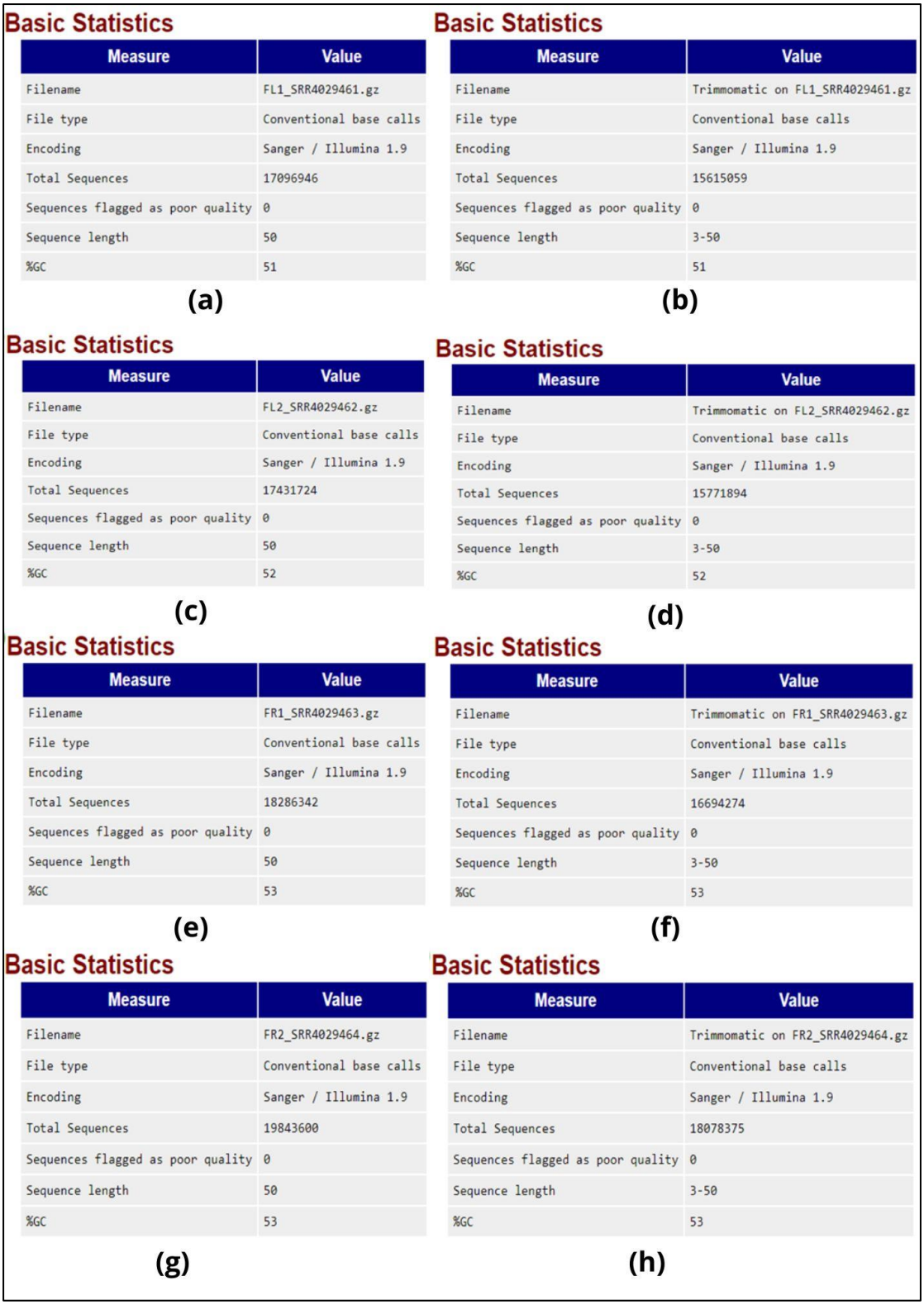
4.1 TRANSCRIPTOME ANALYSIS

In the present study, the datasets (SRR4029464, SRR4029463, SRR4029462 and SRR4029461) from NCBI (https://www.ncbi.nlm.nih.gov/) were used to study the differentially expressed genes in the left and right ovary of chicken during its embryonic 12^{th} day to identify the genes responsible for the regression of right ovary during the embryonic period.

4.1.1 Quality control and trimming of RNA-Seq data

The reads quality was evaluated using the FASTQC tool. The quality of reads was evaluated based on a number of parameters including fundamental statistics like per-sequence quality scores, per-base sequence content, adapter content, per-sequence GC content, sequence length distribution, per-base N content, sequence duplication levels, overrepresented sequences, per-tile sequence quality and K-mer content. FastQC reported sequence counts for each sample are shown in Table 4.1. The mean quality value across each base position in the read of all the samples is 35.0919 as demonstrated in Fig 4.2. The number of reads with average quality scores (Phred quality) is shown in Fig 4.3. The average GC content of reads (Per Sequence GC Content) is shown in Fig 4.4. The unique reads of the samples were mentioned in Fig.4.5. A total of 66.159602 million processed reads were obtained after filtration for its quality (Q30) using Trimmomatic. The FastQC report was found to have satisfactory quality parameters and does not have any adapter sequences.

Quality trimming was performed using the Trimmomatic tool where reads were truncated based on their average quality (30). A total of 34528670 and 38129942 single-end reads of left and right ovary samples, respectively were generated with a read length of 50 bp. After the removal of 3141717 and 3357293 low-quality reads from left and right ovary samples, respectively, the rest of 31386953 (90.9%) and 34772649 (91.2%) reads of left and right ovary samples, were further used for downstream analysis (Table 4.2). To conclude the total raw reads were 72658612 whereas the total processed reads were 66159602 as shown in table 4.1.



Basic Statistics

Measure	Value
Filename	FL1_SRR4029461.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	17096946
Sequences flagged as poor quality	0
Sequence length	50
%GC	51

(a)

Basic Statistics

Measure	Value
Filename	Trimmomatic on FL1_SRR4029461.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	15615059
Sequences flagged as poor quality	0
Sequence length	3-50
%GC	51

(b)

Basic Statistics

Measure	Value
Filename	FL2_SRR4029462.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	17431724
Sequences flagged as poor quality	0
Sequence length	50
%GC	52

(c)

Basic Statistics

Measure	Value
Filename	Trimmomatic on FL2_SRR4029462.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	15771894
Sequences flagged as poor quality	0
Sequence length	3-50
%GC	52

(d)

Basic Statistics

Measure	Value
Filename	FR1_SRR4029463.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	18286342
Sequences flagged as poor quality	0
Sequence length	50
%GC	53

(e)

Basic Statistics

Measure	Value
Filename	Trimmomatic on FR1_SRR4029463.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	16694274
Sequences flagged as poor quality	0
Sequence length	3-50
%GC	53

(f)

Basic Statistics

Measure	Value
Filename	FR2_SRR4029464.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	19843600
Sequences flagged as poor quality	0
Sequence length	50
%GC	53

(g)

Basic Statistics

Measure	Value
Filename	Trimmomatic on FR2_SRR4029464.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	18078375
Sequences flagged as poor quality	0
Sequence length	3-50
%GC	53

(h)

Fig. 4.1 : Basic Statistics of FASTQ files (a) Left Ovary 1 Raw reads (b) Left Ovary 1 Processed reads (c) Left Ovary 2 Raw reads (d) Left Ovary 2 Processed reads (e) Right Ovary 1 Raw reads (f) Right Ovary 1 Processed reads (g) Right Ovary 2 Raw reads (h) Right Ovary 2 Processed reads

Table 4.1 Sequence counts for each sample

S.No	Sample	File Name	Total Raw Reads	Total Processed Reads
1.	Left Ovary 1	FL1_SRR4029461	17096946	15615059
2.	Left Ovary 2	FL2_SRR4029462	17431724	15771894
3.	Right Ovary 1	FR1_SRR4029463	18286342	16694274
4.	Right Ovary 2	FR2_SRR4029464	19843600	18078375
Total			72658612	66159602

Table 4.2 Trimming report of left and right ovary samples

Sample	Input reads	Retained reads	Percentage %
Left Ovary	34528670	31386953	90.9
Right Ovary	38129942	34772649	91.2

4.1.2 Mapping to reference genome

Cleaned and high-quality reads were further used for reference-based alignment. Clean readings from each sample were mapped by HISAT2 to the *Gallus gallus* reference genome (GRCg7b) (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/016/699/485/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b/GCF_016699485.2_bGalGal1.mat.broiler.GRCg7b_genomic.fna.gz). When the processed reads were aligned to the *Gallus gallus* reference genome (GRCg7b), a total of 53.063506 million reads were mapped resulting in overall mapping percentage of 80.22. individual sample wise mapping results were provided in Table 4.3 and Fig 4.6.

4.1.3 Transcript Quantification

Mapped BAM files from HISAT2 are looked for gene annotation with the help of the featureCounts tool, which counts the number of reads per annotated gene. An exon, a gene, transcripts, or binding area can all be considered features. After the quantification, the output from featureCounts is required for differential gene analysis.

A total of 23539837 (68.43%) and 26085009 (69.67%) reads were assigned to annotated genes in left and right ovary samples respectively. Unassigned-Unmapped (UU), Unassigned-Multi-Mapping (UMM), Unassigned-No-Features (UNF), and Unassigned-Ambiguity (UA) were 11058198, 7729061, 1374865, and 2063795 reads respectively, for combined left and right ovary samples.

Table 4.3 Mapping Summary

S.No	Sample	mapped uniquely	multi-mapped	not aligned	Unique Mapping %
1.	Left Ovary 1	12606593	583289	2425177	80.73
2.	Left Ovary 2	12679856	459200	2632838	80.39
3.	Right Ovary 1	13358985	471096	2864193	80.02
4.	Right Ovary 2	14418072	524313	3135990	79.75
Total		53063506	2037898	11058198	80.22

Table 4.4 Transcript Quantification Summary

Sample	Assigned	UU	UMM	UNF	UA	Total	Assigned %
Left Ovary 1	11718719	2425177	2345543	397703	490171	17377313	67.44
Left Ovary 2	11821118	2632838	1718474	369183	489555	17031168	69.41
Left Ovary Total	23539837	5058015	4064017	766886	979726	34408481	68.43
Right Ovary 1	12537700	2864193	1728783	300103	521182	17951961	69.84
Right Ovary 2	13547309	3135990	1936261	307876	562887	19490323	69.50
Right Ovary Total	26085 009	6000183	3665044	607979	1084069	37442284	69.67
Overall Total	49624846	11058198	7729061	1374865	2063795	71850765	69
Note: **UU** - Unassigned: Unmapped; **UMM**- Unassigned: Multi Mapping; **UNF**- Unassigned: No Features; **UA**- Unassigned: Ambiguity							

Percentage reads assigned to the annotated gene in the left ovary sample 1 and left ovary sample 2 were 67.44 % and 69.41% respectively, whereas in the right ovary sample 1

and right ovary sample 2 were 69.84% and 69.50 % respectively. The overall reads assigned to annotated genes in all the samples is 69%. Transcript quantification summary is provided in Table 4.4 and Fig. 4.7.

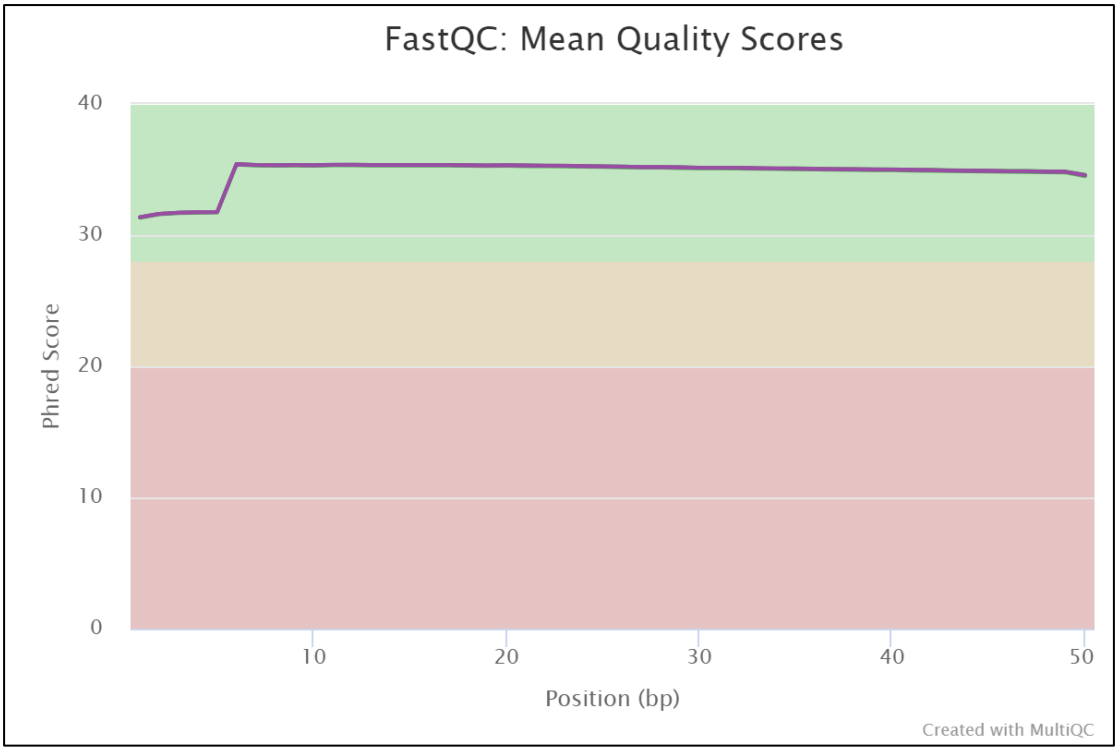


Fig. 4.2 The mean quality value across each base position in the read

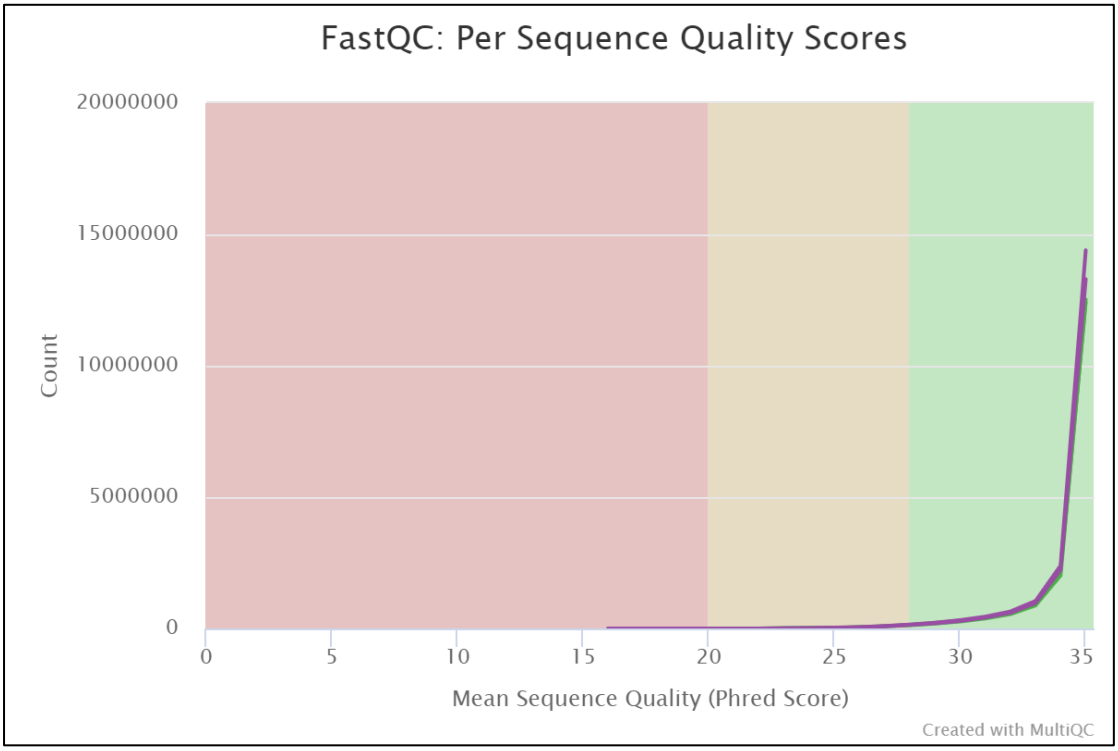


Fig. 4.3 The mean sequence quality (Phred quality) score

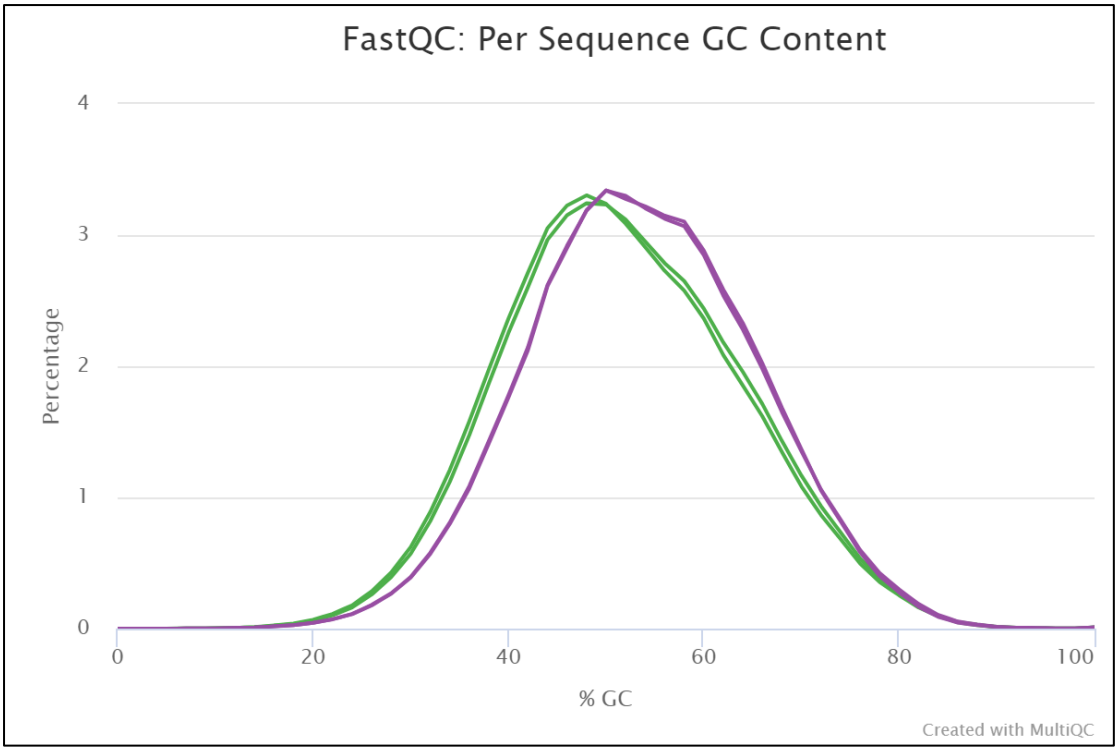


Fig. 4.4 Per Sequence GC Content (Green – Left ovary; Purple - Right Ovary)

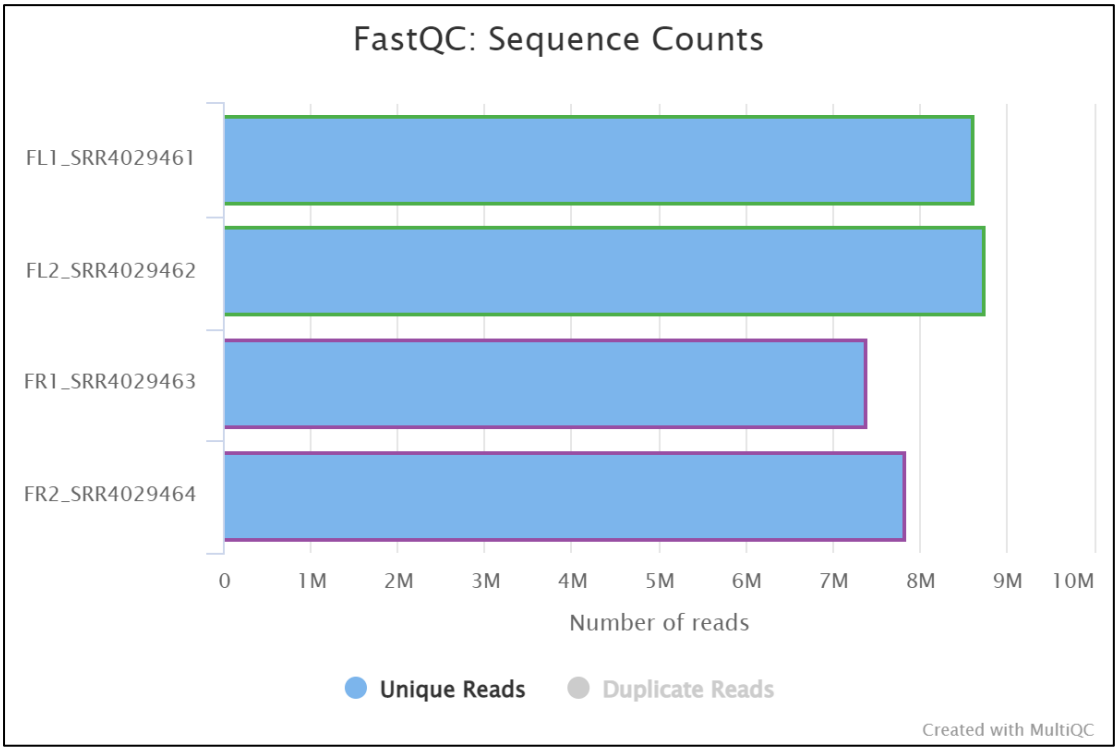


Fig. 4.5 Unique sequence counts (Green – Left ovary; Purple - Right Ovary)

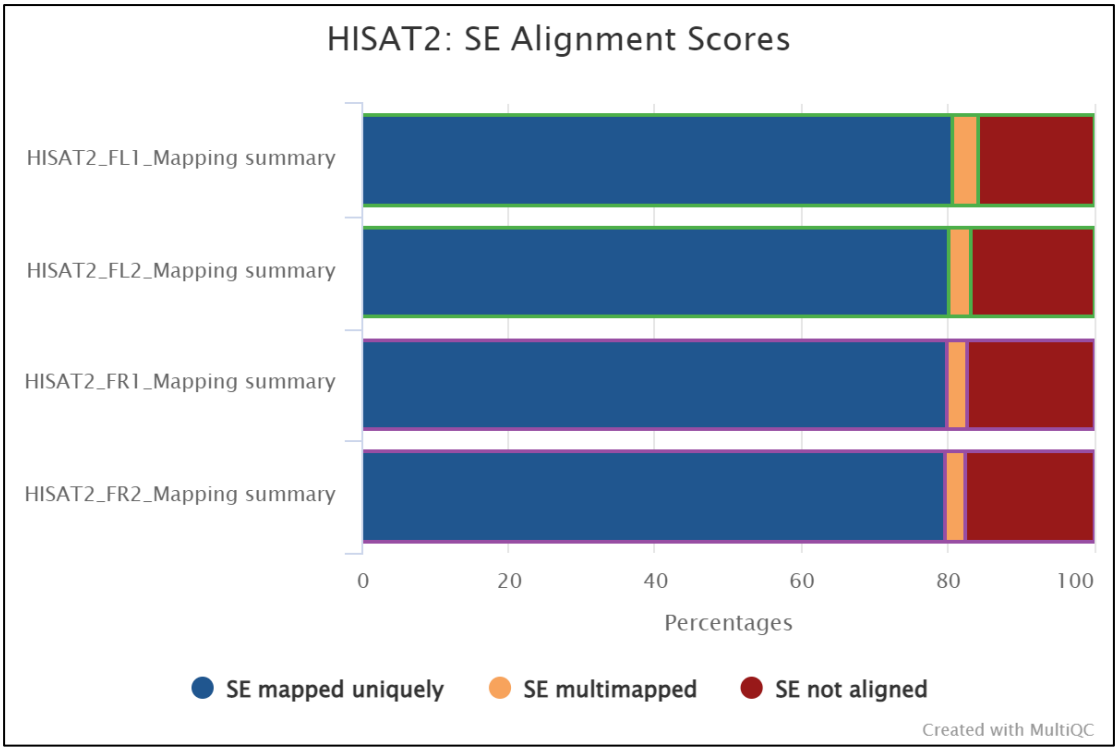


Fig. 4.6 Sequence alignment with the genome

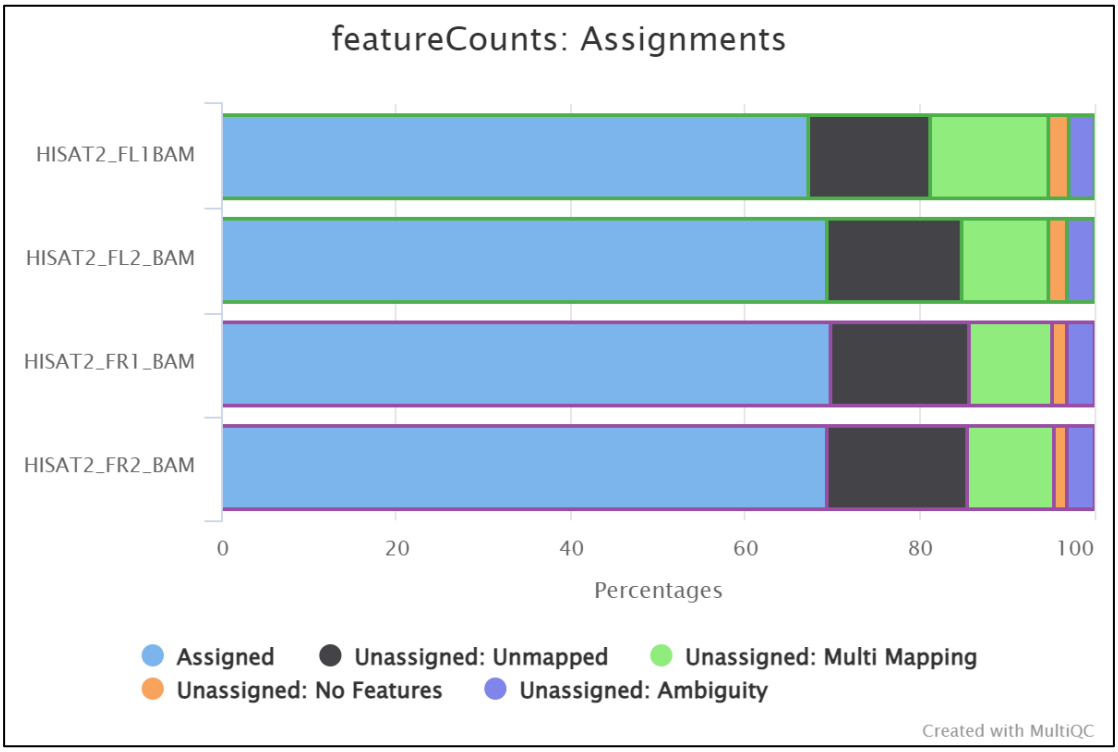


Fig. 4.7 featureCounts summary (Green – Left ovary; Purple - Right Ovary)

4.1.4 Differential genes estimation

Differential genes analysis performed by DESeq2, which was an R-language statistical package that can be used to derive useful data and visual diagrams from statistical analysis of gene readings. Deseq2 generates Principal Component Analysis (PCA) which is a technique used to emphasize variation and bring out strong patterns in a dataset (dimensionality reduction) PCA plot (Fig. 4.8). The representation of log fold-change vs mean expression between left and right ovaries done using MA plot (Figure 4.9 a). Base 2 log fold-change is plotted along the y-axis of a scatter plot, while normalized mean expression plotted along the x-axis. The representation of dispersion versus the mean of normalized counts of left and right ovaries demonstrated in Figure 4.9 (b). Each black dot in the plot represents the dispersion of one gene. The red line is fit to data (black bots), and then the dispersions are squeezed toward the red line, resulting in the final (blue) dispersion estimates. Usually, the dispersion is highest at the low counts and levels off at higher counts. A sample-sample distance heatmap (Figure 4.10) was used to visualize overall gene expression between left and ovary samples (n = 8). The heatmap was constructed using the R package DESeq2.

The total number of significantly up and down-regulated differentially expressed genes between embryonic left and right chicken ovaries during embryonic 12^{th} day were found to be 9152. Similarly, the total number of significantly up-regulated genes by a fold change above +2 in the right ovary during E12 was 479 and the significantly down-regulated genes by fold change below -2 in the right ovary during E12 were 227. A list of the top 100 up-regulated and down-regulated genes in the right ovary are listed in Table 4.5 and 4.6 respectively.

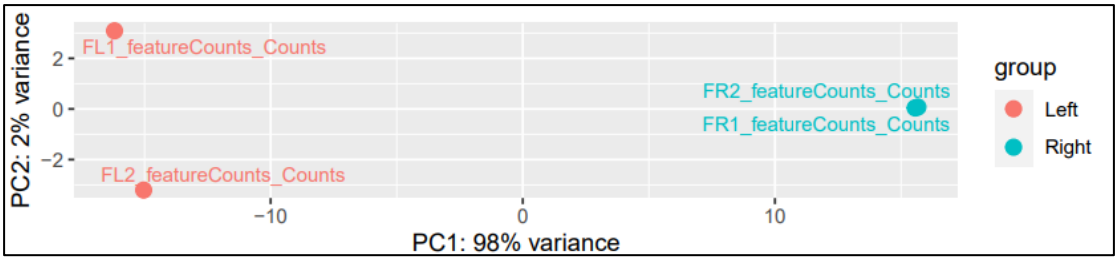
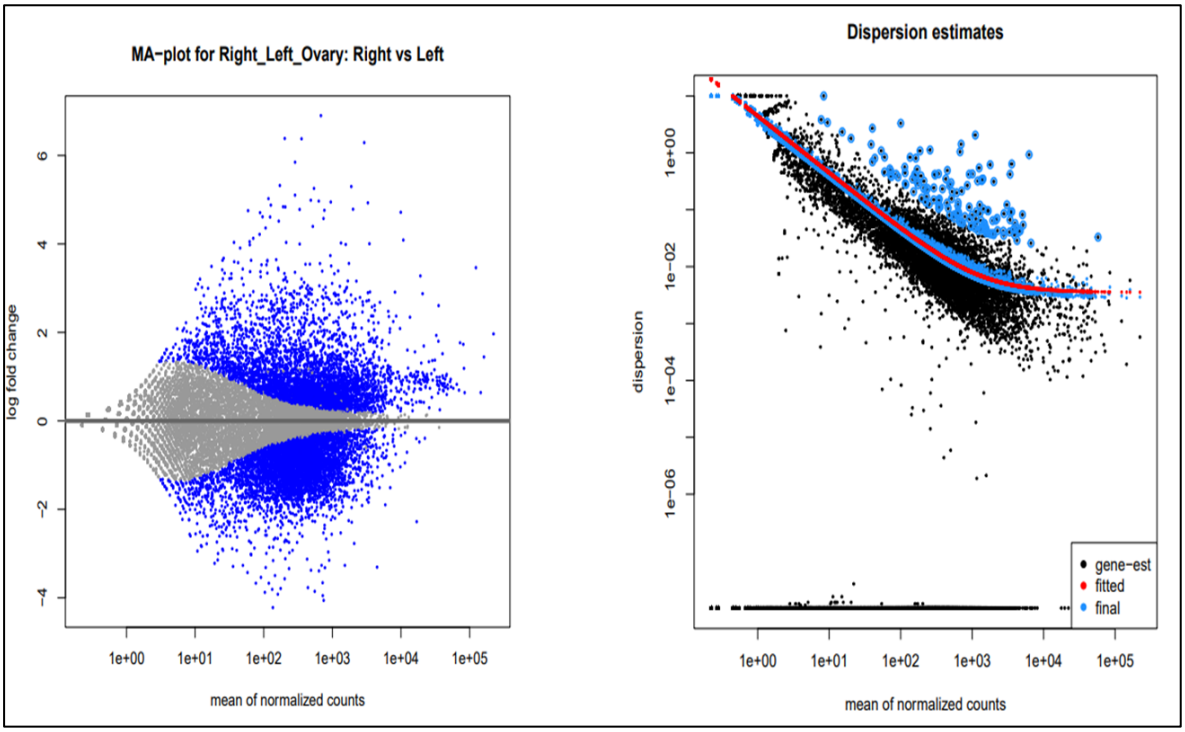


Fig. 4.8 Principal Component Analysis (PCA)



(a) **(b)**

Fig. 4.9 (a) MA-plot (b) Dispersion vs normalised counts

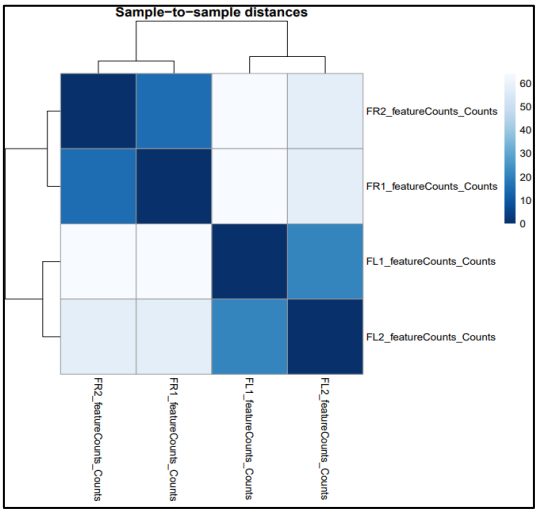


Fig. 4.10 Heatmap of Euclidian distances between left and ovary samples

Table 4.5 List of top 100 up-regulated genes in right ovary with fold change

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
1	ORM1	681.0294142	6.900756711	0.308336112	22.38063091	6.08E-111	2.02E-108
2	C8A	202.3493106	6.38461556	0.449032876	14.21859267	7.03E-46	4.77E-44
3	PIT54	355.5925149	6.378028718	0.356203043	17.90559861	1.07E-71	1.54E-69
4	APOA2	2913.130113	6.291334202	0.150511865	41.7995896	0	0
5	APOV1	286.0733089	5.848798895	0.346044363	16.90187593	4.36E-64	5.37E-62
6	LOC100858647	171.7453795	5.321935686	0.386427667	13.77213936	3.75E-43	2.32E-41
7	SPP2	1887.535814	5.298366126	0.15108503	35.06876964	2.02E-269	4.80E-266
8	RBP4A	540.0188969	5.256182601	0.233883503	22.47350724	7.54E-112	2.61E-109
9	TM4SF4	285.7220263	5.10360731	0.298837626	17.0781952	2.16E-65	2.74E-63
10	CPB2	117.2252229	5.025017329	0.426826634	11.77297041	5.38E-32	1.94E-30
11	APOH	485.3675522	4.988787522	0.411568355	12.12140695	8.14E-34	3.25E-32
12	NME4	633.6398853	4.969308201	0.459876716	10.80573996	3.23E-27	8.73E-26
13	FETUB	948.3228148	4.948995532	0.177898768	27.81916699	2.54E-170	3.26E-167
14	TTR	3292.273388	4.929547971	0.120248615	40.99463406	0	0
15	LOC396477	126.9156591	4.878425473	0.401558314	12.14873483	5.83E-34	2.35E-32
16	AHSG	191.966567	4.840778396	0.337521034	14.34215327	1.19E-46	8.31E-45
17	SPIA4	206.0628402	4.83699376	0.331713196	14.58185509	3.66E-48	2.77E-46
18	GC	2002.091888	4.785053387	0.276594414	17.29989163	4.71E-67	6.23E-65

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
19	SERPINC1	333.2836176	4.780443631	0.261958429	18.24886354	2.11E-74	3.38E-72
20	FGA	765.3789811	4.761104596	0.186468386	25.53303913	8.47E-144	6.72E-141
21	LOC423629	59.12762208	4.752902931	0.529812811	8.970909779	2.94E-19	4.55E-18
22	TF	9925.230223	4.717070806	0.111675738	42.2389937	0	0
23	HRG	88.39121629	4.707818655	0.451598227	10.42479437	1.91E-25	4.59E-24
24	LEAP2	96.02238828	4.614438594	0.438525361	10.52262653	6.80E-26	1.68E-24
25	TMPRSS6	123.7467121	4.609921618	0.38918599	11.84503486	2.28E-32	8.44E-31
26	FGG	742.9510754	4.575874432	0.180680835	25.3257321	1.66E-141	1.20E-138
27	ANGPTL3	165.4873053	4.516715422	0.339050403	13.32166365	1.73E-40	9.71E-39
28	ITIH2	518.9638194	4.396637641	0.203283271	21.62813308	9.77E-104	2.90E-101
29	GCGR	106.4849769	4.361504583	0.407047562	10.71497532	8.66E-27	2.26E-25
30	SERPINA10	87.06693111	4.217819131	0.42657765	9.887576459	4.71E-23	9.51E-22
31	IYD	70.35078052	4.127315467	0.466607483	8.845369224	9.12E-19	1.37E-17
32	ALB	10813.27891	4.089175885	0.117325188	34.8533505	3.79E-266	7.88E-263
33	LOC425662	62.28891931	4.053502313	0.477642415	8.486478971	2.13E-17	2.87E-16
34	AQP3	36.86065705	4.035983855	0.567125625	7.116560563	1.11E-12	9.66E-12
35	FABP1	212.2845884	4.034900345	0.624688475	6.459060002	1.05E-10	7.42E-10
36	AMBP	1335.647821	4.022000599	0.299962581	13.40834108	5.40E-41	3.08E-39
37	APOC3	3467.043178	4.006041178	0.108326456	36.98118938	2.30E-299	6.37E-296
38	SERPIND1	334.0375362	3.99895934	0.231324908	17.28719733	5.87E-67	7.70E-65

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
39	HMGCS2	1566.401484	3.988636415	0.307008929	12.99192317	1.36E-38	7.01E-37
40	PROC	62.65143516	3.981489992	0.465853152	8.546663205	1.27E-17	1.74E-16
41	HNF4beta	45.28158688	3.90812304	0.523377511	7.467120694	8.20E-14	7.93E-13
42	VTN	935.3827418	3.737502215	0.154805862	24.14315684	8.81E-129	4.31E-126
43	LOC428593	26.04864693	3.677626434	0.608622858	6.042537485	1.52E-09	9.40E-09
44	KNG1	151.8844837	3.645188147	0.562098774	6.484960145	8.88E-11	6.33E-10
45	LOC107052201	25.11941951	3.631533655	0.610120632	5.952156775	2.65E-09	1.59E-08
46	MBL2	65.41003214	3.631235679	0.439690824	8.258611458	1.47E-16	1.84E-15
47	RNP	113.2215163	3.629270708	0.368550018	9.84743056	7.03E-23	1.40E-21
48	SLC38A4	112.7911409	3.613312487	0.348985196	10.35377009	4.02E-25	9.35E-24
49	A2ML2	37.14117403	3.598434191	0.542436689	6.633832603	3.27E-11	2.46E-10
50	LOC107053815	42.32262428	3.577912478	0.51338919	6.96920104	3.19E-12	2.64E-11
51	MUC4L	1743.113421	3.559574941	0.126582778	28.12053104	5.50E-174	7.62E-171
52	SELENOP2	659.946269	3.557391542	0.166447002	21.37251803	2.41E-101	6.79E-99
53	PLG	470.593765	3.54423971	0.191191312	18.53766089	1.03E-76	1.71E-74
54	C3	406.1373255	3.541616266	0.225774659	15.68650923	1.87E-55	1.81E-53
55	IDO2	92.79909673	3.528574064	0.377577708	9.345292335	9.16E-21	1.60E-19
56	F7	109.6455649	3.527576551	0.349839712	10.08340801	6.54E-24	1.39E-22
57	SPINK5	961.5497592	3.502698426	0.151203907	23.16539622	1.02E-118	3.94E-116
58	SPIA3	17.34264715	3.492858899	0.663026799	5.268050863	1.38E-07	6.83E-07

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
59	AvBD10	123508.2089	3.464905833	0.079781104	43.43015641	0	0
60	PAH	162.2633465	3.455726605	0.296039904	11.67317839	1.75E-31	6.11E-30
61	TMEM207	49.47290954	3.432957467	0.47797078	7.182358436	6.85E-13	6.08E-12
62	LBFABP	24.8554252	3.415304503	0.603981376	5.654652011	1.56E-08	8.62E-08
63	TMEM72	163.690316	3.41413979	0.296786789	11.50367846	1.26E-30	4.17E-29
64	NPHS2	1214.01722	3.390818433	0.139433188	24.3185893	1.25E-130	6.48E-128
65	PLET1	711.2215713	3.385103012	0.160464438	21.09565873	8.72E-99	2.30E-96
66	RHCG	63.26190302	3.375186602	0.431335252	7.824972778	5.08E-15	5.50E-14
67	F12	89.60990802	3.37426569	0.379021025	8.902581822	5.46E-19	8.30E-18
68	SERPINF2	285.3688267	3.358993889	0.229171314	14.65713067	1.21E-48	9.30E-47
69	LOC422154	112.8559592	3.342620188	0.339436641	9.847552636	7.02E-23	1.40E-21
70	F9	26.90049342	3.340317356	0.595541229	5.608876756	2.04E-08	1.11E-07
71	AGT	200.2816934	3.331883582	0.26675525	12.49041429	8.42E-36	3.73E-34
72	AGR2	33.72351553	3.327361831	0.544714733	6.108448382	1.01E-09	6.35E-09
73	AvBD9	1068.764758	3.326924244	0.140422292	23.69227982	4.33E-124	1.95E-121
74	PTGDS	19171.80533	3.278214534	0.088484512	37.04845568	1.90E-300	6.33E-297
75	FOXD2	90.19807398	3.258225156	0.3700869	8.803946199	1.32E-18	1.96E-17
76	CLDN3	208.6544292	3.243282078	0.261519128	12.40170118	2.56E-35	1.10E-33
77	TDO2	33.81855735	3.225187187	0.53637341	6.012951283	1.82E-09	1.12E-08
78	GRB7	48.1003287	3.222191736	0.474077513	6.796761393	1.07E-11	8.44E-11

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
79	MRGPRH	31.70530217	3.21792699	0.553913369	5.80944092	6.27E-09	3.62E-08
80	IL8L2	25.37044944	3.210904394	0.607551452	5.284991723	1.26E-07	6.27E-07
81	HSD11B1B	354.9222425	3.20848623	0.223099637	14.38140499	6.77E-47	4.77E-45
82	LOC419389	206.0628723	3.207134353	0.26111615	12.2824052	1.13E-34	4.77E-33
83	LOC121106432	11.77761902	3.205977002	0.694983415	4.61302663	3.97E-06	1.60E-05
84	PGLYRP2	193.5717339	3.182973319	0.265195102	12.0023835	3.45E-33	1.32E-31
85	HPD	1678.271619	3.163270842	0.134310641	23.55190049	1.20E-122	5.26E-120
86	CL2	2238.747154	3.15947467	0.359656611	8.784697885	1.57E-18	2.31E-17
87	TSPAN1	218.6939197	3.150619872	0.249330408	12.63632423	1.33E-36	6.24E-35
88	F2	387.3138049	3.128582986	0.203697168	15.35899109	3.08E-53	2.74E-51
89	HGD	991.8576406	3.108809657	0.136649419	22.75025883	1.43E-114	5.16E-112
90	LOC112530658	11.0196761	3.096915552	0.698181205	4.435690234	9.18E-06	3.51E-05
91	SLC23A3	126.0916439	3.083418689	0.316401627	9.745268125	1.93E-22	3.72E-21
92	SLCO1B3	50.6217496	3.080162066	0.460396137	6.690243075	2.23E-11	1.70E-10
93	TFAP2E	23.0099714	3.076666148	0.614219654	5.009064963	5.47E-07	2.50E-06
94	BAAT	437.2624547	3.058799856	0.200562643	15.25109468	1.62E-52	1.40E-50
95	TTC36	646.8753948	3.054624108	0.168631024	18.11424749	2.46E-73	3.76E-71
96	XPNPEP2	1285.854426	3.042231759	0.130862765	23.2474971	1.51E-119	6.27E-117
97	APOA4	361.4971635	3.036315023	0.209251457	14.510365	1.04E-47	7.71E-46
98	GPIHBP1	163.4834713	3.027792239	0.283196434	10.69149139	1.12E-26	2.88E-25

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
99	AADAC	795.4887946	3.024797934	0.161933935	18.67920974	7.31E-78	1.25E-75
100	SULT	1627.77887	3.020067891	0.14390816	20.98607808	8.79E-98	2.25E-95

Table 4.6 List of top 100 down-regulated genes in right ovary with fold-change

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
1	RGS7	135.8820672	-4.220065803	0.369970329	-11.40649797	3.88E-30	1.24E-28
2	DDX4	747.1710023	-4.056071105	0.18625172	-21.77736187	3.80E-105	1.19E-102
3	STK31	725.0026375	-3.949854572	0.18912114	-20.88531496	7.28E-97	1.84E-94
4	BOLL	160.4618381	-3.918302023	0.350239346	-11.1874981	4.69E-29	1.41E-27
5	LOC776720	103.9039467	-3.875985502	0.389552339	-9.949845271	2.53E-23	5.22E-22
6	TDRD9	553.307011	-3.853007709	0.213548442	-18.04278068	8.99E-73	1.34E-70
7	SLC1A3	133.0131858	-3.807627804	0.353002899	-10.78639246	3.99E-27	1.07E-25
8	RBM46	176.4259644	-3.788381299	0.335386515	-11.29556833	1.38E-29	4.26E-28
9	ELAVL2	440.4990399	-3.768090096	0.223082023	-16.8910522	5.24E-64	6.41E-62
10	SERPINB2	118.6483726	-3.669486975	0.370681143	-9.899308455	4.19E-23	8.49E-22
11	TDRD12	76.8310348	-3.668462045	0.438414364	-8.367568102	5.88E-17	7.61E-16
12	DAZL	721.6450469	-3.658263555	0.207219669	-17.65403629	9.47E-70	1.35E-67
13	SMC1B	201.8883628	-3.568235959	0.319087997	-11.18260792	4.96E-29	1.48E-27

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
14	MOV10L1	91.11013275	-3.53670651	0.397915526	-8.888083719	6.22E-19	9.41E-18
15	DDX43	145.9521243	-3.495579789	0.323499927	-10.8055041	3.24E-27	8.73E-26
16	MYO3B	58.06387207	-3.426690733	0.462982991	-7.401331798	1.35E-13	1.28E-12
17	RNF17	637.5880121	-3.390585666	0.411316307	-8.243256117	1.68E-16	2.08E-15
18	TMED11	488.2924273	-3.375450622	0.193354501	-17.457316	3.03E-68	4.13E-66
19	ZPLD1	233.8049897	-3.338516608	0.262258129	-12.72988802	4.03E-37	1.95E-35
20	MX1	779.5921183	-3.317947439	0.17428856	-19.03709255	8.41E-81	1.55E-78
21	TDRD1	106.8624493	-3.314313927	0.367087627	-9.028672412	1.74E-19	2.74E-18
22	LOC422926	4482.231091	-3.305810097	0.097516498	-33.9000084	6.66E-252	1.23E-248
23	MYH13	35.69463996	-3.289474687	0.549169189	-5.989911216	2.10E-09	1.28E-08
24	TRPM3	49.68286984	-3.266090929	0.483998606	-6.748141192	1.50E-11	1.16E-10
25	TUBA3E	1522.707855	-3.265827212	0.12499498	-26.1276669	1.77E-150	1.55E-147
26	TOM1L1	696.6874897	-3.265326455	0.162627578	-20.0785531	1.14E-89	2.49E-87
27	LOC121108334	19.22873915	-3.218937278	0.646650269	-4.977864281	6.43E-07	2.91E-06
28	TEX14	137.0706326	-3.156720971	0.319482036	-9.880746392	5.05E-23	1.02E-21
29	GREB1	897.6236928	-3.149604651	0.172561147	-18.25210781	1.99E-74	3.22E-72
30	MSMB	21.3472624	-3.143106965	0.631093301	-4.98041567	6.34E-07	2.88E-06
31	RASD2	429.0725218	-3.12492268	0.196975959	-15.8644877	1.12E-56	1.13E-54
32	SYCP3	206.2890639	-3.11504643	0.27701667	-11.24497825	2.45E-29	7.50E-28
33	LOC121108482	65.6116509	-3.024552717	0.431428549	-7.010553026	2.37E-12	1.99E-11

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
34	LOC124417990	64.85530623	-3.013688373	0.457089082	-6.593218899	4.30E-11	3.18E-10
35	LOC768589	21.58964203	-2.997809726	0.616090125	-4.865862322	1.14E-06	4.96E-06
36	GREM2	109.2390205	-2.994487276	0.343620702	-8.714513584	2.92E-18	4.19E-17
37	PTPRZ1	236.4925507	-2.979089216	0.252326638	-11.80647925	3.61E-32	1.32E-30
38	TRIM2	870.6728771	-2.963902263	0.171697815	-17.26231786	9.04E-67	1.18E-64
39	RIPPLY3	10.81536474	-2.941809518	0.699755297	-4.204054661	2.62E-05	9.35E-05
40	OVALY	25.50092808	-2.927746165	0.586169373	-4.994710235	5.89E-07	2.68E-06
41	GAD1	40.39680319	-2.926814443	0.51298121	-5.705500289	1.16E-08	6.51E-08
42	ASZ1	263.0067332	-2.925965405	0.247837352	-11.80599043	3.63E-32	1.33E-30
43	MOBP	174.9219179	-2.920159125	0.281894723	-10.35904145	3.81E-25	8.88E-24
44	LOC424300	25.06444605	-2.897669468	0.587342671	-4.933524522	8.08E-07	3.60E-06
45	LOC101749146	34.7716103	-2.888646421	0.533360578	-5.415935376	6.10E-08	3.16E-07
46	USP13	73.17362037	-2.874396244	0.397284314	-7.235111333	4.65E-13	4.20E-12
47	TERB1	120.5327288	-2.857424192	0.36054372	-7.925319551	2.28E-15	2.54E-14
48	SHC4	136.7997375	-2.854259055	0.30991785	-9.20972786	3.27E-20	5.46E-19
49	SYCP2	12.42321449	-2.838339236	0.685385283	-4.141231666	3.45E-05	0.000120765
50	LOC121108996	12.20366682	-2.81883603	0.685934956	-4.109480072	3.97E-05	0.000137275
51	LOC112531728	14.74342225	-2.776954842	0.67006136	-4.144329175	3.41E-05	0.000119279
52	HTR1B	2080.137297	-2.769919011	0.109378406	-25.32418515	1.73E-141	1.20E-138
53	LOC101750051	65.01133189	-2.753298215	0.42835775	-6.427567185	1.30E-10	9.02E-10

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
54	SPON2	306.5609643	-2.747594576	0.218783453	-12.55851183	3.57E-36	1.62E-34
55	ATF3	233.1314562	-2.743821001	0.239555854	-11.45378398	2.25E-30	7.33E-29
56	LOC101747264	20.76861361	-2.737657757	0.618340558	-4.427427119	9.54E-06	3.63E-05
57	PLN	200.8284386	-2.733021116	0.271944741	-10.04991347	9.19E-24	1.95E-22
58	ADARB2	61.1733709	-2.715773297	0.432445228	-6.280039922	3.38E-10	2.25E-09
59	DGKQ	217.6992742	-2.715289545	0.264230993	-10.27619626	9.02E-25	2.06E-23
60	CLGN	207.3768638	-2.691020504	0.283592407	-9.489042851	2.33E-21	4.18E-20
61	CA3B	140.8862759	-2.67675667	0.305868012	-8.751345568	2.11E-18	3.08E-17
62	PIWIL1	813.2105425	-2.668837801	0.156420334	-17.06196209	2.85E-65	3.59E-63
63	STK17B	159.3995959	-2.657672585	0.289183216	-9.190272586	3.92E-20	6.50E-19
64	FMN2	146.8952814	-2.647650746	0.297224418	-8.907918008	5.20E-19	7.92E-18
65	LOC121110329	40.87823843	-2.623105562	0.486966543	-5.386623787	7.18E-08	3.68E-07
66	LOC101749477	364.3993482	-2.606998218	0.200995486	-12.97043166	1.80E-38	9.22E-37
67	LOC121110769	12.9547153	-2.606372835	0.676771176	-3.851187707	0.000117546	0.000373652
68	LOC107052149	22.63731477	-2.603937122	0.590474276	-4.409907811	1.03E-05	3.92E-05
69	MYF5	8.557545809	-2.599900233	0.70691721	-3.677800167	0.000235254	0.00071231
70	EML5	32.5394619	-2.599524415	0.530708425	-4.898215841	9.67E-07	4.26E-06
71	CCDC169	28.66341909	-2.594355689	0.5519095	-4.700690397	2.59E-06	1.08E-05
72	LOC121106818	15.04583519	-2.586550074	0.661171937	-3.912068749	9.15E-05	0.000296873
73	MEIOC	15.22120804	-2.585465893	0.663736078	-3.895322219	9.81E-05	0.000316627

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
74	ESR1	110.2663451	-2.582521359	0.333658363	-7.740016866	9.94E-15	1.05E-13
75	TDRD15	105.9746936	-2.581573313	0.376820535	-6.85093585	7.34E-12	5.85E-11
76	CPNE4	177.3223948	-2.573406959	0.276813576	-9.296534482	1.45E-20	2.48E-19
77	NPY4R	8.333380197	-2.563456353	0.707425459	-3.623641643	0.000290484	0.000865214
78	ZC3HAV1	978.0724458	-2.562837614	0.137708364	-18.610617	2.64E-77	4.48E-75
79	RNF212	10.17924939	-2.554496938	0.695428925	-3.673268174	0.000239468	0.000723752
80	DNAH7	228.3883013	-2.550532931	0.241202074	-10.57425787	3.92E-26	9.85E-25
81	BRCA1	334.3501683	-2.544419202	0.219763671	-11.57797917	5.33E-31	1.79E-29
82	PTBP3	1323.413678	-2.53174297	0.156029114	-16.22609336	3.30E-59	3.64E-57
83	LOC121106668	25.63115762	-2.531731203	0.572275193	-4.423975097	9.69E-06	3.68E-05
84	MAEL	740.6752242	-2.530563624	0.158989458	-15.91654979	4.86E-57	5.00E-55
85	PARPBP	75.70338771	-2.513965705	0.381443883	-6.590656754	4.38E-11	3.23E-10
86	LOC101747704	11.81049319	-2.512666247	0.679834779	-3.695995446	0.000219027	0.000666572
87	RGS22	24.53981661	-2.511609661	0.595066901	-4.220718136	2.44E-05	8.72E-05
88	LOC121111441	38.23534352	-2.496783066	0.498690681	-5.006676809	5.54E-07	2.53E-06
89	WDR41	106.7421537	-2.485453737	0.356966315	-6.962712266	3.34E-12	2.76E-11
90	LOC107053365	82.70292675	-2.473144602	0.364678264	-6.781716503	1.19E-11	9.31E-11
91	TPH2	56.82414181	-2.463068946	0.457883528	-5.379247767	7.48E-08	3.82E-07
92	CFAP99	27.98513358	-2.455434708	0.552418615	-4.444880462	8.79E-06	3.37E-05
93	ARC	84.90669384	-2.442095919	0.373599469	-6.536668595	6.29E-11	4.56E-10

Sl No	GENE ID	Base Mean	Log2FC	Std err	Wald-stats	P value	P-adj
94	LOC107051715	769.2098837	-2.431868428	0.153318069	-15.86159054	1.17E-56	1.17E-54
95	RGPD1	20.40489454	-2.430261331	0.604371151	-4.021140533	5.79E-05	0.000195611
96	GPR149	177.4348194	-2.415464988	0.265298543	-9.10470505	8.65E-20	1.40E-18
97	LOC101747677	118.0228585	-2.411900626	0.312263359	-7.723930953	1.13E-14	1.18E-13
98	PKP1	44.76084831	-2.409042806	0.466530445	-5.163741904	2.42E-07	1.16E-06
99	LOC769098	36.73181804	-2.407605295	0.522235326	-4.610192332	4.02E-06	1.62E-05
100	BNC1	285.9735415	-2.407241564	0.234334808	-10.27265897	9.36E-25	2.13E-23

4.1.5 Volcano Plot

A scatterplot that displays statistical significance (P value) vs magnitude of change (fold change) known as a volcano plot. It makes it possible to quickly visually identify genes that have significant statistical fold changes. These genes might be the ones with the most biological impact. A Volcano Plot (Fig. 4.11) was generated showing up and down-regulated genes in the right ovary with a significance threshold of 0.05 and LogFC threshold to colour as 2.0. Only the top most significant 50 genes were labelled in the plot. The red colour indicated regulated genes and the blue colour indicates down-regulated genes.

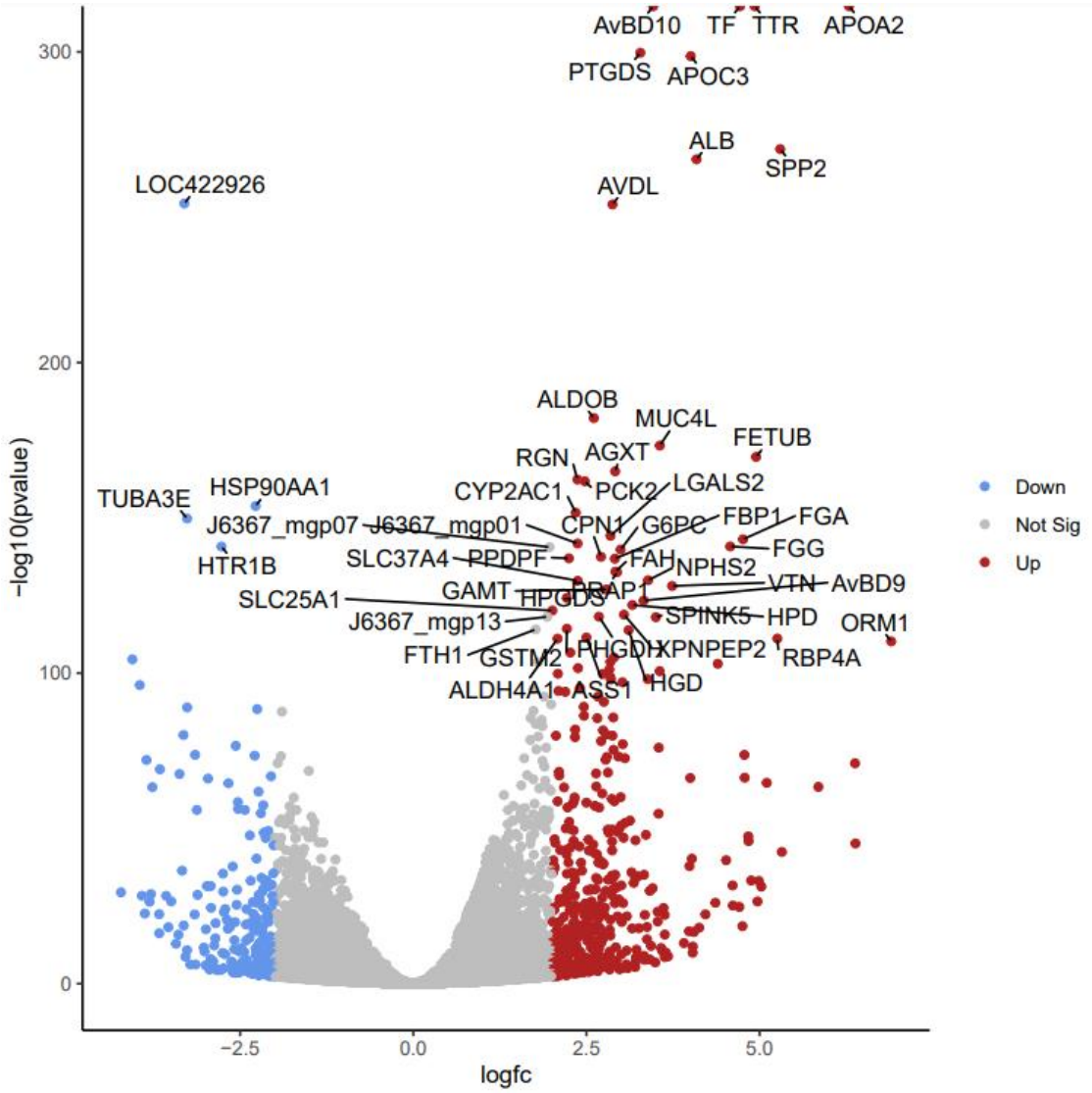


Fig. 4.11 Volcano Plot

4.1.6 Heatmap

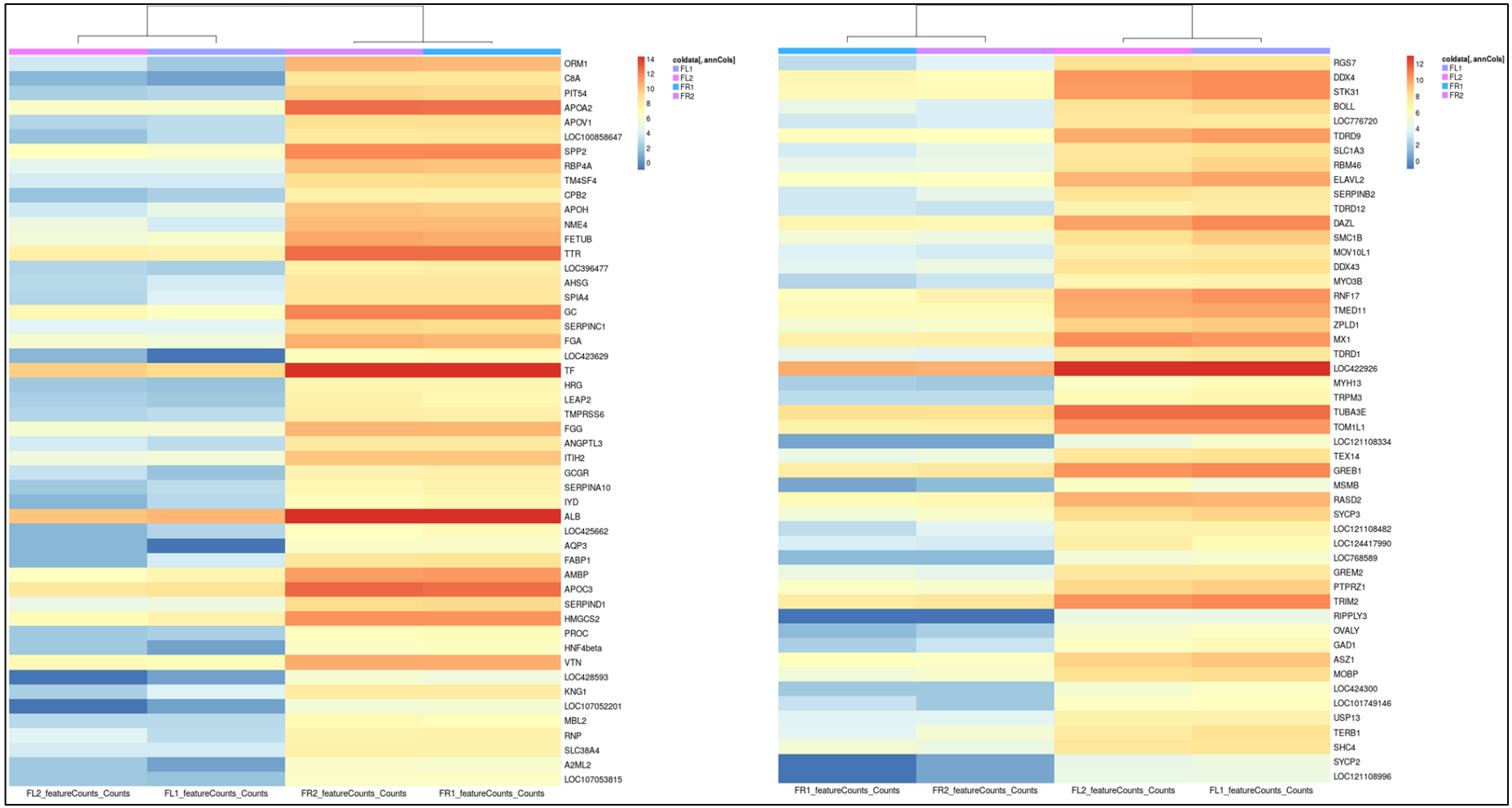
A common method of visualising gene expression data is to display it as a heatmap. Identifying genes that are differentially expressed among left and right ovaries can be visually analysed using a heatmap. In heat maps, the data is shown as a grid with each column denoting an ovary sample and each row denoting a gene. Changes in gene expression are represented by the colour and intensity of the boxes. In the figure 4.12, orange to red represents up-regulated genes, whereas blue represents down-regulated genes.

Heatmap of top 50 up-regulated genes in the right ovary (ORM1, C8A, PIT54, APOA2, APOV1, LOC100858647, SPP2, RBP4A, TM4SF4, CPB2, APOH, NME4, FETUB, TTR, LOC396477, AHSG, SPIA4, GC, SERPINC1, FGA, LOC423629, TF, HRG, LEAP2, TMPRSS6, FGG, ANGPTL3, ITIH2, GCGR, SERPINA10, IYD, ALB, LOC425662, AQP3, FABP1, AMBP, APOC3, SERPIND1, HMGCS2, PROC, HNF4beta, VTN, LOC428593, KNG1, LOC107052201, MBL2, RNP, SLC38A4, A2ML2 and LOC107053815) were shown in Figure 4.12a.

Heatmap of top 50 down-regulated genes in the right ovary (RGS7, DDX4, STK31, BOLL, LOC776720, TDRD9, SLC1A3, RBM46, ELAVL2, SERPINB2, TDRD12, DAZL, SMC1B, MOV10L1, DDX43, MYO3B, RNF17, TMED11, ZPLD1, MX1, TDRD1, LOC422926, MYH13, TRPM3, TUBA3E, TOM1L1, LOC121108334, TEX14, GREB1, MSMB, RASD2, SYCP3, LOC121108482, LOC124417990, LOC768589, GREM2, PTPRZ1, TRIM2, RIPPLY3, OVALY, GAD1, ASZ1, MOBP, LOC424300, LOC101749146, USP13, TERB1, SHC4, SYCP2 and LOC121108996) were shown in Figure 4.12b.

4.1.7 Gene Ontology enrichment and pathway analysis of DEGs

Gene ontology enrichment analysis to identify functional processes of biological importance within the list of differentially expressed genes was performed using the gProfiler and ShinyGO. All differentially expressed genes were classified into the categories of molecular function, cellular component, and biological process.



(a) **(b)**

Fig 4.12 Heatmap of (a) Up-regulated genes in Right ovary (b) Down-regulated genes in Right ovary

4.1.7.1 Enrichment analysis of upregulated genes

The top 50 up-regulated genes with the highest fold change (Figure 4.12a) were further processed for pathway analysis. Enrichment annotations of 39 genes were found, and details of the genes are included in Table 4.7. Enrichment analysis by g:Profiler for up-regulated genes in right ovary mentioned in Table 4.8. Pathway Analysis of up-regulated genes in the right ovary by ShinyGO 0.76 was performed and the pathway tree plot, Pathway-network and enrichment bar plot are shown in Figures 4.15 a, 4.15 b and 4.16 respectively. Enrichment process of up-regulated genes in right analysed from STRING database (Figure 4.13).

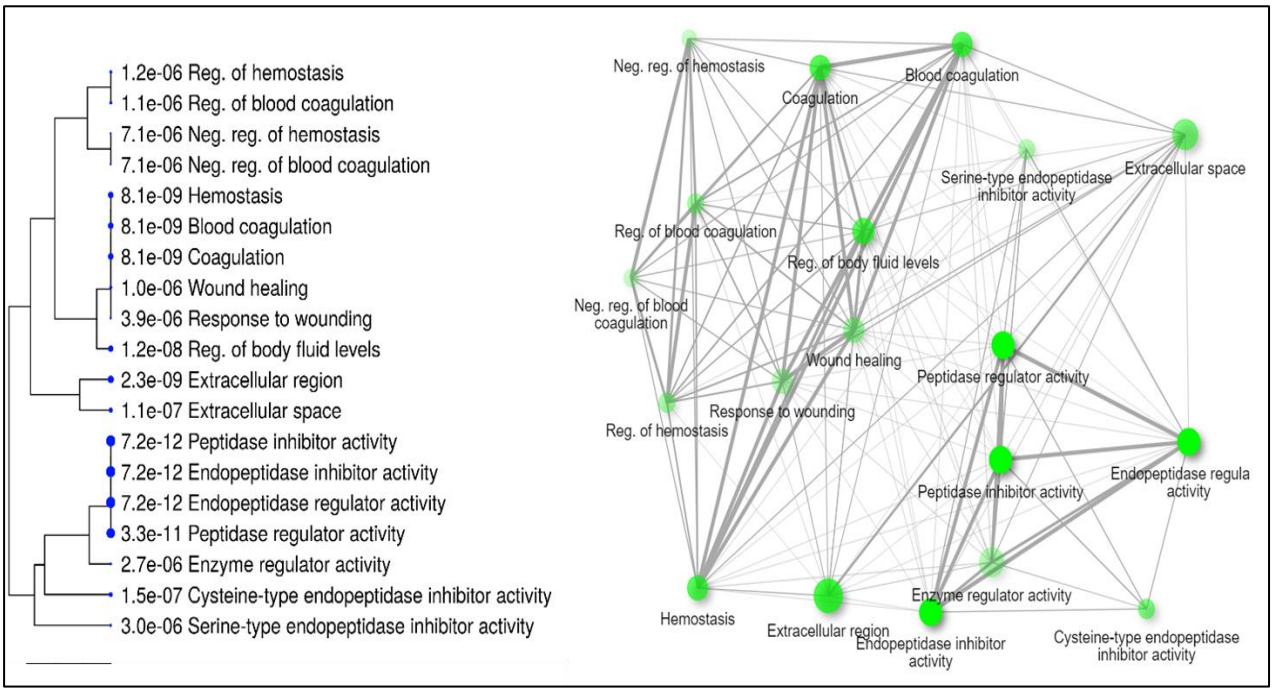
Sl. No	Term id	Genes	Enrichment process
1	GO:0045861	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,FABP1,CPB2,FETUB	Negative regulation of proteolysis
2	GO:0010951	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,FABP1,FETUB	Negative regulation of endopeptidase activity
3	GO:0051346	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,ANGPTL3,FABP1,FETUB	Negative regulation of hydrolase activity
4	GO:0007596	PROC,SERPINC1,HRG,FGA,FGG,SERPINA10,CPB2	Blood coagulation
5	GO:0050878	PROC,AQP3,SERPINC1,HRG,FGA,FGG,SERPINA10,CPB2	Regulation of body fluid levels
6	GO:0030193	PROC,SERPINC1,HRG,CPB2	Regulation of blood coagulation
7	GO:0065008	PROC,AQP3,SERPINC1,TF,HRG,FGA,FGG,SERPINA10,ANGPTL3,GIPR,IYD,TMPRSS6,TTR,CPB2,RNP	Regulation of biological quality
8	GO:0006950	PROC,AQP3,SERPINC1,TF,AHSG,HRG,FGA,FGG,C8A,SERPINA10,GIPR,FABP1,CPB2	Response to stress
9	GO:0030195	PROC,HRG,CPB2	Negative regulation of blood coagulation
10	GO:0030168	HRG,FGA,FGG	Platelet activation
11	GO:0051918	HRG,CPB2	Negative regulation of fibrinolysis
12	GO:0072378	FGA,FGG	Blood coagulation, fibrin clot formation

Fig. 4.13 Enrichment process of up-regulated genes in right ovary from STRING database

4.1.7.2 Enrichment analysis of down-regulated genes

The top 50 down-regulated genes with the highest fold change (Figure 4.12b) were further processed for pathway analysis. Enrichment annotations of 35 genes were found, and details of the genes are included in Table 4.9. Enrichment analysis by g:Profiler for up-regulated genes in right ovary mentioned in Table 4.10.

Fig. 4.14 gProfiler Pathway Analysis of up-regulated genes in right ovary



(a) **(b)**

Fig. 4.15 Pathway Analysis of up-regulated genes in right ovary (a) Tree plot (b) Pathway-network

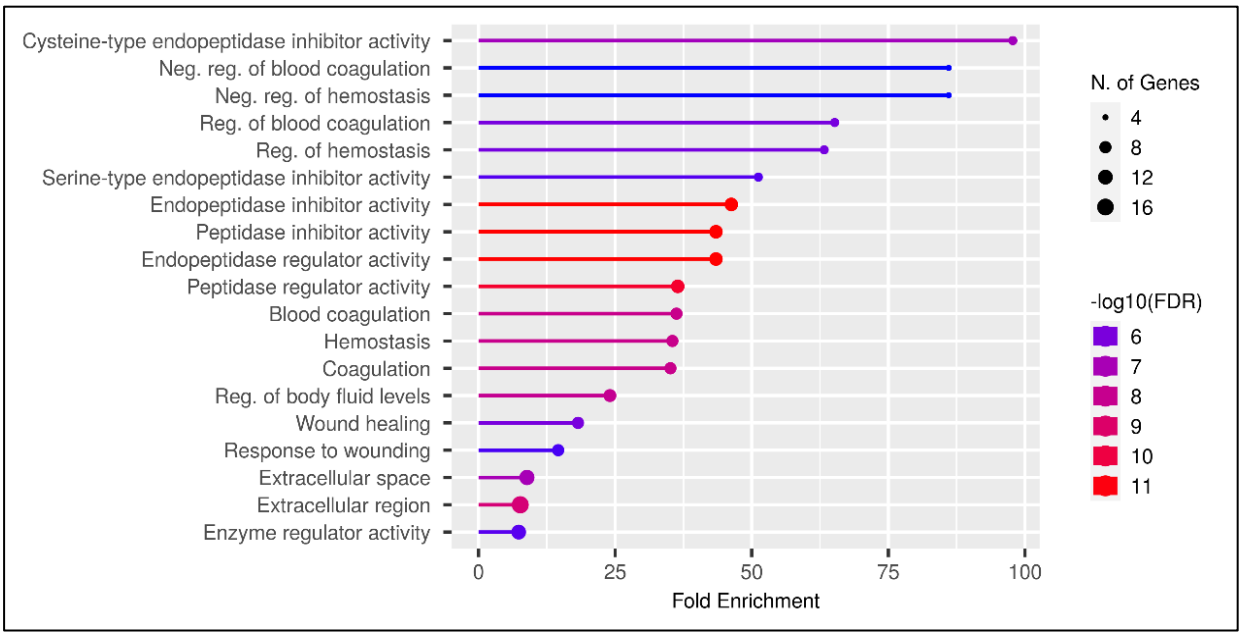


Fig 4.16 ShinyGO 0.76 enrichment bar-plot of up-regulated genes in right ovary

Table 4.7 Details of up-regulated genes in right ovary

Sl No	Gene Symbol	Ensemble Gene ID	Chr	Description
1	ITIH2	ENSGALG00000006833	1	inter-alpha-trypsin inhibitor heavy chain 2 [Source:NCBI gene;Acc:419110]
2	SLC38A4	ENSGALG00000009730	1	solute carrier family 38 member 4 [Source:NCBI gene;Acc:417809]
3	TMPRSS6	ENSGALG00000012483	1	transmembrane protease, serine 6 [Source:NCBI gene;Acc:418046]
4	CPB2	ENSGALG00000016985	1	carboxypeptidase B2 [Source:NCBI gene;Acc:418851]
5	TTR	ENSGALG00000015143	2	transthyretin [Source:NCBI gene;Acc:396277]
6	IYD	ENSGALG00000012392	3	iodotyrosine deiodinase [Source:NCBI gene;Acc:421631]
7	FGA	ENSGALG00000009266	4	fibrinogen alpha chain [Source:NCBI gene;Acc:396307]
8	FGG	ENSGALG00000009268	4	fibrinogen gamma chain [Source:NCBI gene;Acc:395837]
9	GC	ENSGALG00000011612	4	GC, vitamin D binding protein [Source:NCBI gene;Acc:395696]
10	ALB	ENSGALG00000035219	4	albumin [Source:NCBI gene;Acc:396197]
11	FABP1	ENSGALG00000015937	4	fatty acid binding protein 1 [Source:NCBI gene;Acc:374015]
12	SERPINA10	ENSGALG00000020391	5	serpin family A member 10 [Source:NCBI gene;Acc:423432]
13	SPIA4	ENSGALG00000010969	5	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4 [Source:NCBI gene;Acc:423433]
14	LOC423629	ENSGALG00000050806	6	uncharacterized LOC423629 [Source:NCBI gene;Acc:423629]
15	LOC100858647	ENSGALG00000044811	6	beta-microseminoprotein-like [Source:NCBI gene;Acc:100858647]
16	SPP2	ENSGALG00000004129	7	secreted phosphoprotein 2 [Source:NCBI gene;Acc:396080]

Sl No	Gene Symbol	Ensemble Gene ID	Chr	Description
17	HMGCS2	ENSGALG00000002960	8	3-hydroxy-3-methylglutaryl-CoA synthase 2 [Source:NCBI gene;Acc:424380]
18	SERPINC1	ENSGALG00000004591	8	serpin peptidase inhibitor, clade C (antithrombin), member 1 [Source:NCBI gene;Acc:424440]
19	C8A	ENSGALG00000010850	8	complement C8 alpha chain [Source:NCBI gene;Acc:429111]
20	ANGPTL3	ENSGALG00000010978	8	angiopoietin like 3 [Source:NCBI gene;Acc:100189558]
21	PROC	ENSGALG00000001780	9	protein C, inactivator of coagulation factors Va and VIIIa [Source:NCBI gene;Acc:395085]
22	TF	ENSGALG00000006453	9	transferrin (ovotransferrin) [Source:NCBI gene;Acc:396241]
23	AHSG	ENSGALG00000008601	9	alpha 2-HS glycoprotein [Source:NCBI gene;Acc:424956]
24	FETUB	ENSGALG00000008608	9	fetuin B [Source:NCBI gene;Acc:395404]
25	HRG	ENSGALG00000008646	9	histidine rich glycoprotein [Source:NCBI gene;Acc:770588]
26	KNG1	ENSGALG00000046159	9	kininogen 1 [Source:NCBI gene;Acc:424957]
27	TM4SF4	ENSGALG00000010427	9	transmembrane 4 L six family member 4 [Source:NCBI gene;Acc:771806]
28	HNF4beta	ENSGALG00000005708	11	hepatic nuclear factor 4beta [Source:NCBI gene;Acc:415820]
29	LEAP2	ENSGALG00000048876	13	liver enriched antimicrobial peptide 2 [Source:NCBI gene;Acc:414338]
30	NME4	ENSGALG00000002688	14	NME/NM23 nucleoside diphosphate kinase 4 [Source:NCBI gene;Acc:395915]
31	SERPIND1	ENSGALG00000039239	15	serpin peptidase inhibitor, clade D (heparin cofactor), member 1 [Source:NCBI gene;Acc:395877]
32	AMBP	ENSGALG00000030255	17	alpha-1-microglobulin/bikunin precursor [Source:NCBI gene;Acc:770795]
33	APOH	ENSGALG00000033376	18	apolipoprotein H [Source:NCBI gene;Acc:417431]

Sl No	Gene Symbol	Ensemble Gene ID	Chr	Description
34	GCGR	ENSGALG00000011219	18	glucagon receptor [Source:NCBI gene;Acc:425670]
35	VTN	ENSGALG00000003589	19	vitronectin [Source:NCBI gene;Acc:395935]
36	RNP	ENSGALG00000028991	21	renal natriuretic peptide [Source:NCBI gene;Acc:100526659]
37	APOC3	ENSGALG00000030920	24	apolipoprotein C3 [Source:NCBI gene;Acc:100859721]
38	PIT54	ENSGALG00000046217	31	PIT54 protein [Source:NCBI gene;Acc:395364]
39	AQP3	ENSGALG00000002452	Z	aquaporin 3 (Gill blood group) [Source:NCBI gene;Acc:426894]

Table 4.8 g:Profiler enrichment analysis for up-regulated genes in right ovary

Sl No	Source	GO Term Name	Term Id	Intersections
1	GO:MF	Endopeptidase inhibitor activity	GO:0004866	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1
2	GO:MF	Enzyme inhibitor activity	GO:0004857	APOV1,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SE
3	GO:MF	Peptidase inhibitor activity	GO:0030414	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1

Sl No	Source	GO Term Name	Term Id	Intersections
4	GO:MF	Endopeptidase regulator activity	GO:0061135	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1
5	GO:MF	Peptidase regulator activity	GO:0061134	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1
6	GO:MF	Serine-type endopeptidase inhibitor activity	GO:0004867	SPIA4,SERPINC1,ITIH2,SERPINA10,AMBP,SERPIND1
7	GO:MF	Enzyme regulator activity	GO:0030234	APOV1,APOH,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1
8	GO:MF	Cysteine-type endopeptidase inhibitor activity	GO:0004869	FETUB,AHSG,HRG,KNG1
9	GO:MF	Molecular function regulator activity	GO:0098772	APOV1,APOH,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,RNP
10	GO:BP	Blood coagulation	GO:0007596	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
11	GO:BP	Hemostasis	GO:0007599	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
12	GO:BP	Coagulation	GO:0050817	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
13	GO:BP	Regulation of body fluid levels	GO:0050878	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,AQP3,SERPIND1,PROC,KNG1

Sl No	Source	GO Term Name	Term Id	Intersections
14	GO:BP	Regulation of blood coagulation	GO:0030193	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
15	GO:BP	Regulation of hemostasis	GO:1900046	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
16	GO:BP	Regulation of coagulation	GO:0050818	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
17	GO:BP	Negative regulation of hemostasis	GO:1900047	CPB2,APOH,FGG,PROC,KNG1
18	GO:BP	Negative regulation of blood coagulation	GO:0030195	CPB2,APOH,FGG,PROC,KNG1
19	GO:BP	Wound healing	GO:0042060	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
20	GO:BP	Negative regulation of coagulation	GO:0050819	CPB2,APOH,FGG,PROC,KNG1
21	GO:BP	Response to wounding	GO:0009611	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
22	GO:BP	Regulation of wound healing	GO:0061041	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
23	GO:BP	Negative regulation of wound healing	GO:0061045	CPB2,APOH,FGG,PROC,KNG1
24	GO:BP	Regulation of response to wounding	GO:1903034	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
25	GO:BP	Negative regulation of response to wounding	GO:1903035	CPB2,APOH,FGG,PROC,KNG1
26	GO:BP	Fibrinolysis	GO:0042730	CPB2,APOH,FGG

Sl No	Source	GO Term Name	Term Id	Intersections
27	GO:BP	Blood coagulation, fibrin clot formation	GO:0072378	APOH,SERPINC1,FGG
28	GO:BP	Protein activation cascade	GO:0072376	APOH,SERPINC1,FGG
29	GO:BP	Regulation of biological quality	GO:0065008	RBP4A,CPB2,APOH,TTR,SERPINC1,FGA,TMPRSS6,FGG,ANGPTL3,GCGR,SERPINA10,IYD,AQP3,SERPIND1,PROC,KNG1,MBL2
30	GO:BP	Response to external stimulus	GO:0009605	RBP4A,CPB2,APOH,AHSG,SERPINC1,LEAP2,FGG,GCGR,ALB,AQP3,PROC,KNG1,MBL2
31	GO:BP	Response to stress	GO:0006950	CPB2,APOH,AHSG,SERPINC1,FGA,LEAP2,FGG,GCGR,SERPINA10,ALB,AQP3,FABP1,SERPIND1,PROC,KNG1,MBL2
32	GO:CC	Extracellular region	GO:0005576	ORM1,APOV1,SPP2,RBP4A,APOH,AHSG,SPIA4,GC,SERPINC1,FGA,TF,TMPRSS6,FGG,ANGPTL3,SERPINA10,ALB,APOC3,SERPIND1,PROC,KNG1,MBL2,RNP
33	GO:CC	Extracellular space	GO:0005615	ORM1,APOV1,RBP4A,APOH,AHSG,SPIA4,GC,SERPINC1,FGA,TF,TMPRSS6,FGG,ANGPTL3,SERPINA10,ALB,SERPIND1,KNG1,MBL2
34	GO:CC	Chylomicron	GO:0042627	APOV1,APOH
35	GO:CC	Fibrinogen complex	GO:0005577	FGA,FGG
36	GO:CC	Very-low-density lipoprotein particle	GO:0034361	APOV1,APOH

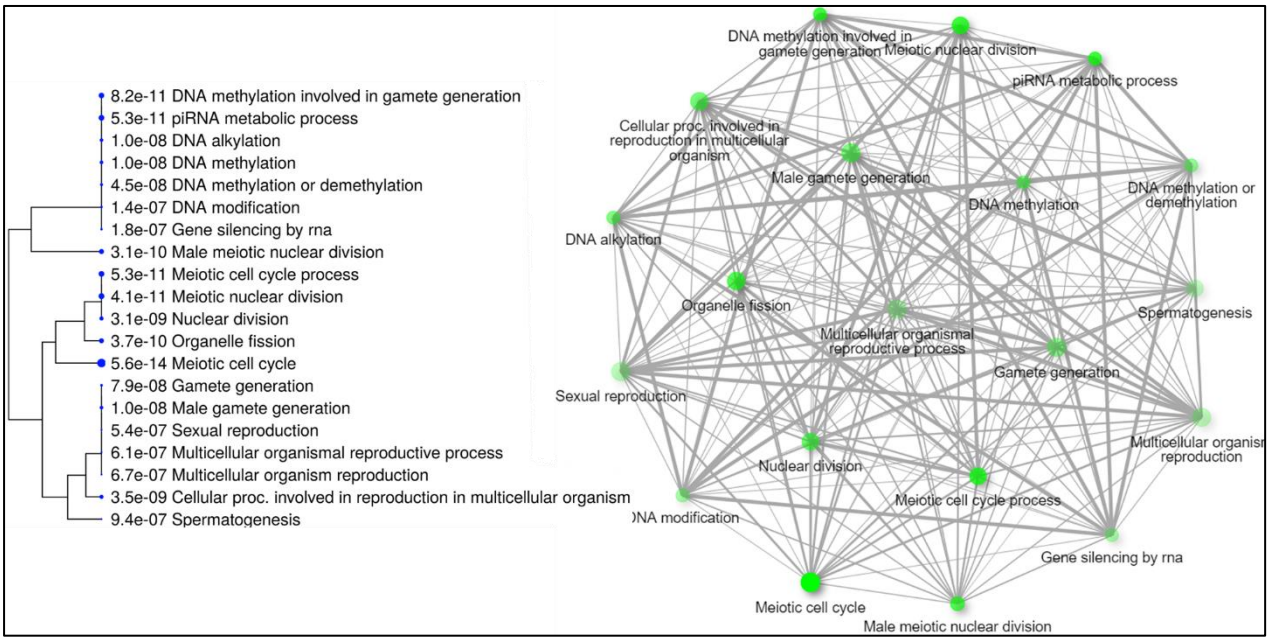
Sl No	Source	GO Term Name	Term Id	Intersections
37	GO:CC	Triglyceride-rich plasma lipoprotein particle	GO:0034385	APOV1,APOH
38	KEGG	PPAR signaling pathway	KEGG:03320	FABP1,APOC3,HMGCS2
39	REAC	Post-translational protein phosphorylation	REAC:R-GGA-8957275	SPP2,AHSG,TF,FGG,ITIH2,SERPINA10,SERPIND1,PROC,KNG1
40	REAC	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (igfbps)	REAC:R-GGA-381426	SPP2,AHSG,TF,FGG,ITIH2,SERPINA10,SERPIND1,PROC,KNG1
41	REAC	Platelet degranulation	REAC:R-GGA-114608	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1
42	REAC	Response to elevated platelet cytosolic Ca2+	REAC:R-GGA-76005	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1
43	REAC	Intrinsic Pathway of Fibrin Clot Formation	REAC:R-GGA-140837	SPIA4,SERPIND1,PROC,KNG1
44	REAC	Formation of Fibrin Clot (Clotting Cascade)	REAC:R-GGA-140877	SPIA4,SERPIND1,PROC,KNG1
45	REAC	Platelet activation, signaling and aggregation	REAC:R-GGA-76002	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1

Sl No	Source	GO Term Name	Term Id	Intersections
46	REAC	Hemostasis	REAC:R-GGA-109582	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,SERPIND1,PROC,KNG1
47	REAC	Common Pathway of Fibrin Clot Formation	REAC:R-GGA-140875	SPIA4,SERPIND1,PROC
48	WP	Blood clotting cascade	WP:WP775	FGA,FGG
49	HP	Recurrent thromboembolism	HP:0004831	SERPINC1,HRG
Notes: BP - Biological Process, MF - Molecular Function, CC - Cellular Component, KEGG - Kyoto Encyclopedia of Genes and Genomes, REAC - Reactome, WP – WikiPathways, HP - human phenotype ontology				

Pathway Analysis of down-regulated genes in the right ovary by ShinyGO 0.76 was performed and the pathway tree plot, Pathway-network and enrichment bar plot are shown in Figure 4.18 a, 4.18 b and 4.19 respectively. The enrichment process of up-regulated genes in right was analysed from the STRING database (Figure 4.17) and found only one enrichment process (Meiotic cell cycle) involved by TDRD12, CCDC79, SYCP3, SMC1B, DDX4 is reported.

Sl No	Term id	Genes	Enrichment process
1	GO:0051321	TDRD12,CCDC79,SYCP3,SMC1B,DDX4	Meiotic cell cycle

Fig. 4.17 Enrichment process of down-regulated genes in the right ovary from STRING database



(a) **(b)**

Fig. 4.18 Pathway Analysis of down-regulated genes in right ovary (a) Tree plot (b) Pathway-network

Fig. 4.19 ShinyGO 0.76 enrichment bar-plot of down-regulated genes in right ovary

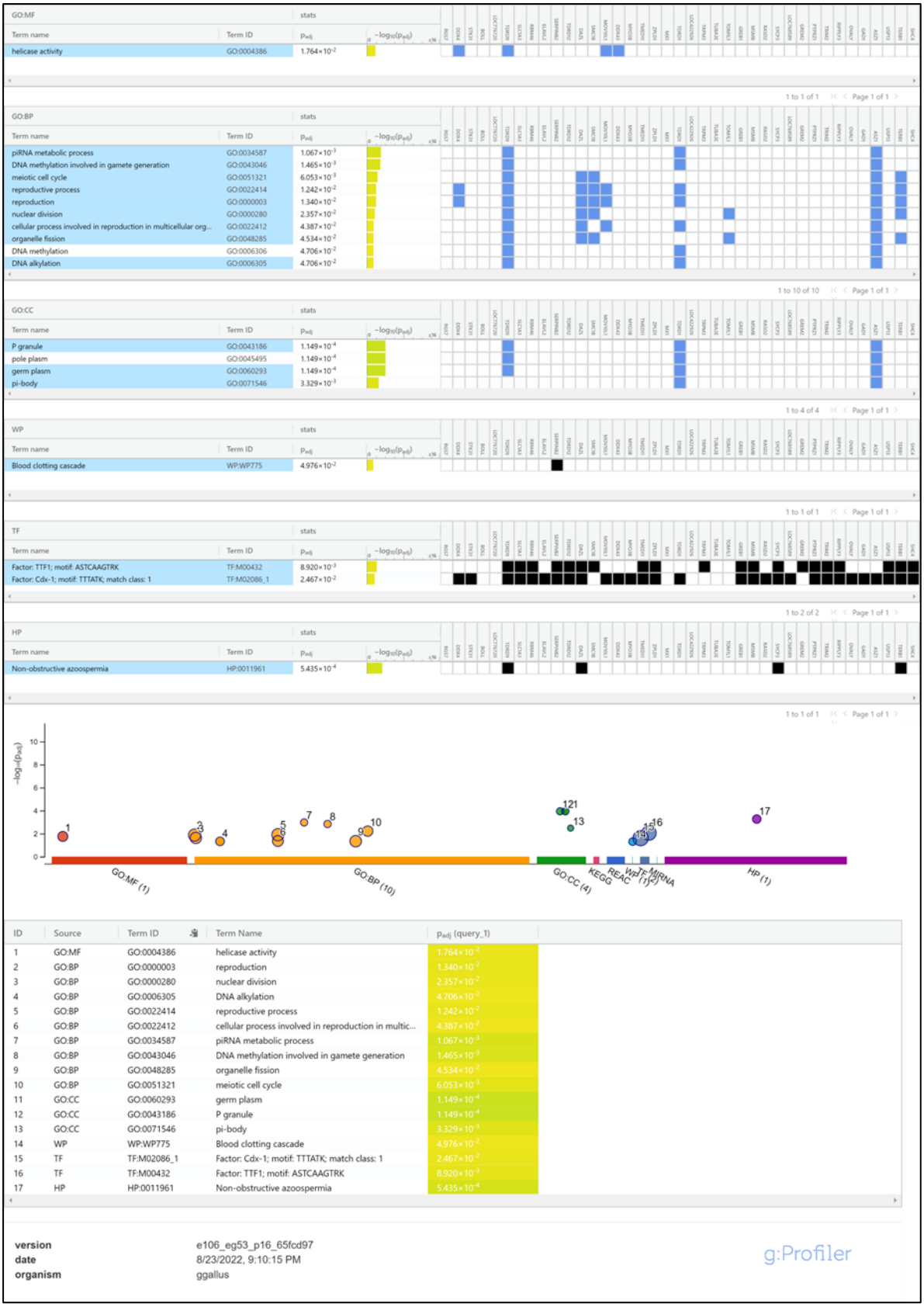


Fig. 4.20 gProfiler Pathway Analysis of down-regulated genes in right ovary

Table 4.9 Details of down-regulated genes in right ovary

Sl No	Symbol	Ensembl Gene ID	Chr	Description
1	MOV10L1	ENSGALG00000033664	1	Mov10 RISC complex RNA helicase like 1 [Source:NCBI gene;Acc:417738]
2	PTPRZ1	ENSGALG00000040114	1	Protein tyrosine phosphatase, receptor type Z1 [Source:NCBI gene;Acc:396403]
3	RASD2	ENSGALG00000012542	1	RASD family member 2 [Source:NCBI gene;Acc:418057]
4	SYCP3	ENSGALG00000012766	1	Synaptonemal complex protein 3 [Source:NCBI gene;Acc:418097]
5	SMC1B	ENSGALG00000014228	1	Structural maintenance of chromosomes 1B [Source:NCBI gene;Acc:418243]
6	ZPLD1	ENSGALG00000015347	1	Zona pellucida like domain containing 1 [Source:NCBI gene;Acc:418405]
7	RIPPLY3	ENSGALG00000054080	1	Ripply transcriptional repressor 3 [Source:NCBI gene;Acc:101749102]
8	Mx	ENSGALG00000016142	1	Myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) [Source:NCBI gene;Acc:395313]
9	STK31	ENSGALG00000010980	2	Serine/threonine kinase 31 [Source:NCBI gene;Acc:420621]
10	DAZL	ENSGALG00000011243	2	Deleted in azoospermia like [Source:NCBI gene;Acc:374054]
11	SERPINB2	ENSGALG00000019552	2	Serpin family B member 2 [Source:NCBI gene;Acc:420896]
12	OVALY	ENSGALG00000019551	2	Ovalbumin-related protein Y (SERPINB14B) [Source:NCBI gene;Acc:420897]
13	RGS7	ENSGALG00000039772	3	Regulator of G-protein signaling 7 [Source:NCBI gene;Acc:395620]
14	GREM2	ENSGALG00000055021	3	Gremlin 2, DAN family BMP antagonist [Source:NCBI gene;Acc:421507]
15	DDX43	ENSGALG00000015928	3	DEAD-box helicase 43 [Source:NCBI gene;Acc:428637]
16	GREB1	ENSGALG00000016455	3	Growth regulating estrogen receptor binding 1 [Source:NCBI gene;Acc:421944]
17	TRIM2	ENSGALG00000009207	4	Tripartite motif containing 2 [Source:NCBI gene;Acc:100857562]

Sl No	Symbol	Ensembl Gene ID	Chr	Description
18	RBM46	ENSGALG00000009278	4	RNA binding motif protein 46 [Source:NCBI gene;Acc:422404]
19	TMED11	ENSGALG00000015722	4	Transmembrane emp24 protein transport domain containing 11 [Source:NCBI gene;Acc:422906]
20	LOC422926	ENSGALG00000027090	4	Uncharacterized LOC422926 [Source:NCBI gene;Acc:422926]
21	MSMB	ENSGALG00000043909	6	Microseminoprotein, beta- [Source:NCBI gene;Acc:101750594]
22	TDRD1	ENSGALG00000046180	6	Tudor domain containing 1 [Source:NCBI gene;Acc:423906]
23	LOC768589	ENSGALG00000037678	7	Baculoviral IAP repeat-containing protein 5.1-like [Source:NCBI gene;Acc:768589]
24	BOLL	ENSGALG00000036160	7	Boule homolog, RNA binding protein [Source:NCBI gene;Acc:424061]
25	MYO3B	ENSGALG00000009602	7	Myosin IIIB [Source:NCBI gene;Acc:424157]
26	USP13	ENSGALG00000008909	9	Ubiquitin specific peptidase 13 (isopeptidase T-3) [Source:NCBI gene;Acc:429286]
27	SHC4	ENSGALG00000005011	10	SHC adaptor protein 4 [Source:NCBI gene;Acc:426482]
28	TDRD12	ENSGALG00000004704	11	Tudor domain containing 12 [Source:NCBI gene;Acc:768954]
29	TERB1	ENSGALG00000005214	11	Telomere repeat binding bouquet formation protein 1 [Source:NCBI gene;Acc:769532]
30	TOM1L1	ENSGALG00000003011	18	Target of myb1 like 1 membrane trafficking protein [Source:NCBI gene;Acc:417391]
31	TUBA3E	ENSGALG00000040586	26	Tubulin, alpha 3e [Source:NCBI gene;Acc:421169]
32	SLC1A3	ENSGALG00000003582	Z	Solute carrier family 1 member 3 [Source:NCBI gene;Acc:395443]
33	DDX4	ENSGALG00000014713	Z	DEAD-box helicase 4 [Source:NCBI gene;Acc:395447]
34	TRPM3	ENSGALG00000015126	Z	Transient receptor potential cation channel subfamily M member 3 [Source:NCBI gene;Acc:427249]
35	ELAVL2	ENSGALG00000015799	Z	ELAV like RNA binding protein 2 [Source:NCBI gene;Acc:770158]

Table 4.10 g:Profiler enrichment analysis for down-regulated genes in right ovary

Sl No	Source	GO Term Name	Term Id	Intersections
1	GO:MF	Helicase activity	GO:0004386	DDX4,TDRD9,MOV10L1,DDX43
2	GO:BP	Pirna metabolic process	GO:0034587	TDRD9,TDRD1,ASZ1
3	GO:BP	DNA methylation involved in gamete generation	GO:0043046	TDRD9,TDRD1,ASZ1
4	GO:BP	Meiotic cell cycle	GO:0051321	TDRD9,DAZL,SMC1B,ASZ1,TERB1
5	GO:BP	Reproductive process	GO:0022414	DDX4,TDRD9,DAZL,SMC1B,MOV10L1,TDRD1,ASZ1,TERB1
6	GO:BP	Reproduction	GO:0000003	DDX4,TDRD9,DAZL,SMC1B,MOV10L1,TDRD1,ASZ1,TERB1
7	GO:BP	Nuclear division	GO:0000280	TDRD9,DAZL,SMC1B,TOM1L1,ASZ1,TERB1
8	GO:BP	Cellular process involved in reproduction in multicellular organism	GO:0022412	TDRD9,DAZL,MOV10L1,TDRD1,ASZ1
9	GO:BP	Organelle fission	GO:0048285	TDRD9,DAZL,SMC1B,TOM1L1,ASZ1,TERB1
10	GO:BP	DNA methylation	GO:0006306	TDRD9,TDRD1,ASZ1
11	GO:BP	DNA alkylation	GO:0006305	TDRD9,TDRD1,ASZ1
12	GO:CC	P granule	GO:0043186	TDRD9,TDRD1,ASZ1
13	GO:CC	Pole plasm	GO:0045495	TDRD9,TDRD1,ASZ1
14	GO:CC	Germ plasm	GO:0060293	TDRD9,TDRD1,ASZ1
15	GO:CC	Pi-body	GO:0071546	TDRD1,ASZ1

Sl No	Source	GO Term Name	Term Id	Intersections
16	WP	Blood clotting cascade	WP:WP775	SERPINB2
19	HP	Non-obstructive azoospermia	HP:0011961	TDRD9,DAZL,SYCP3,TERB1
Notes: BP - Biological Process, MF - Molecular Function, CC - Cellular Component, KEGG - Kyoto Encyclopedia of Genes and Genomes, REAC - Reactome, WP – WikiPathways, HP - human phenotype ontology				

4.1.7.3 Combined Enrichment analysis of up and down-regulated genes

The top 50 down-regulated and 50 up-regulated genes with the highest fold change (Figure 4.10) were combined and further processed for pathway analysis. Enrichment analysis was performed using g:Profiler and ShinyGO. ShinyGO provided the visual appearance of enrichment analysis as a bar-plot (Figure 4.22), tree-plot(Figure 4.24), Network pathway (Figure 4.23) and enrichment process from the STRING database (Figure 4.21). The detailed Enrichment analysis by ShinyGO is provided in Table 4.11. Similar results of enrichment were also seen when analysed with g:Profiler. The enrichment analysis by g:Profiler is mentioned in detail in Table 4.12. FGA and FGG are up-regulated genes in the right ovary, whereas SERPINB2 is down-regulated which involve in the Blood Clotting Cascade pathway (Fig 4.25). For a better understanding of GO linkages have been analysed of GO-terms using the Ancestor chart process of terms as shown in Fig 4.31, 4.32, 4.33, 4.34, 4.35 and the process is discussed in section V.

Sl No	Term Id	Gene Names	Enrichment Process
1	GO:0010951	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,SERPINB2,OVALY,FABP1,FETUB	Negative regulation of endopeptidase activity
2	GO:0045861	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,SERPINB2,OVALY,FABP1,CPB2,FETUB	Negative regulation of proteolysis
3	GO:0051346	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,ANGPTL3,SERPINB2,OVALY,FABP1,FETUB	Negative regulation of hydrolase activity
4	GO:0009611	PROC,SLC1A3,SERPINC1,HRG,FGA,FGG,SERPINA10,SERPINB2,CPB2	Response to wounding
5	GO:0007596	PROC,SERPINC1,HRG,FGA,FGG,SERPINA10,CPB2	Blood coagulation
6	GO:0042060	PROC,SERPINC1,HRG,FGA,FGG,SERPINA10,SERPINB2,CPB2	Wound healing
7	GO:0050878	PROC,AQP3,SERPINC1,HRG,FGA,FGG,SERPINA10,CPB2	Regulation of body fluid levels
8	GO:0032269	SERPINC1,ITIH2,AHSG,HRG,SERPINA10,SERPINA4,RASD2,SERPINB2,OVALY,FABP1,CPB2,FETUB	Negative regulation of cellular protein metabolic process
9	GO:0030193	PROC,SERPINC1,HRG,CPB2	Regulation of blood coagulation
10	GO:0010605	SERPINC1,TDRD12,ITIH2,AHSG,HRG,RBM46,SERPINA10,SERPINA4,TMPRSS6,RASD2,SERPINB2,OVALY,FABP1,CPB2,DDX4,FETUB,DAZL	Negative regulation of macromolecule metabolic process
11	GO:0051172	SERPINC1,ITIH2,AHSG,HRG,RBM46,SERPINA10,SERPINA4,TMPRSS6,RASD2,SERPINB2,OVALY,FABP1,CPB2,FETUB,DAZL	Negative regulation of nitrogen compound metabolic process

Fig. 4.21 Enrichment process of up and down regulated genes in right ovary from STRING database

In the present study the up-regulated genes in the right ovary reported in the literature so far are APOH, AHSG, FGG, SERPINA10, SERPIND1, TSPAN1, TSPAN8, SLC26A6, HINTW, PAX2, PITX3, GAL, TNNI1, SPATA2L, CYP11A1. The down-regulated genes reported in the literatures so far are PITX2, DMRT1, DND1, G0S2, SOX9, BMP7, IGF1, CYP19A1, CYP19A1, TDRD5, FSHR, PIWIL1, GREB1, SLC1A3, STK31. Fig 4.36 demonstrates fold changes of the above-mentioned genes.

4.1.7.3.1 KEGG Pathways

g:Profiler analysis of DEGs revealed KEGG pathways involved in right ovary degeneration. FABP1, APOC3, HMGCS2 are up-regulated genes in the right ovary involved in the PPAR signalling pathway (KEGG:03320), demonstrated in Fig 4.26. GCGR, LOC428593, and KNG1 are involved in Neuroactive ligand-receptor interaction (KEGG:04080) as shown in Fig 4.27. TF is up-regulated gene involved in Ferroptosis (KEGG:04216) Fig 4.29, whereas SHC4 is a down-regulated gene involved in the Insulin signalling pathway (KEGG:04910) Fig 4.28. LOC768589, and TUBA3E are down-regulated genes involved in Apoptosis (KEGG:04210) as shown in Fig 4.30.

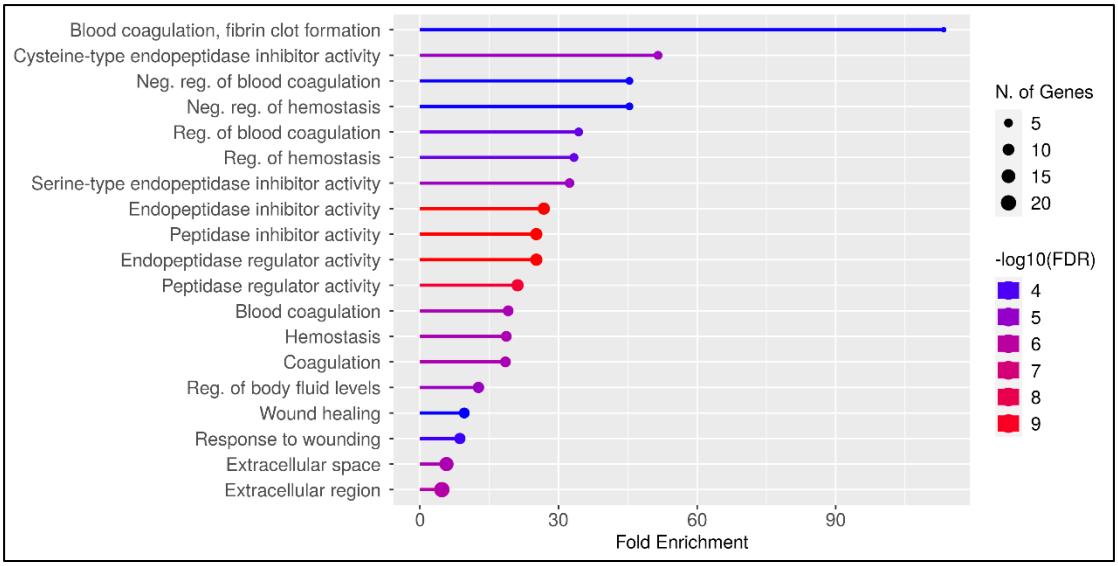


Fig. 4.22 ShinyGO enrichment bar-plot of up and down-regulated genes in right ovary

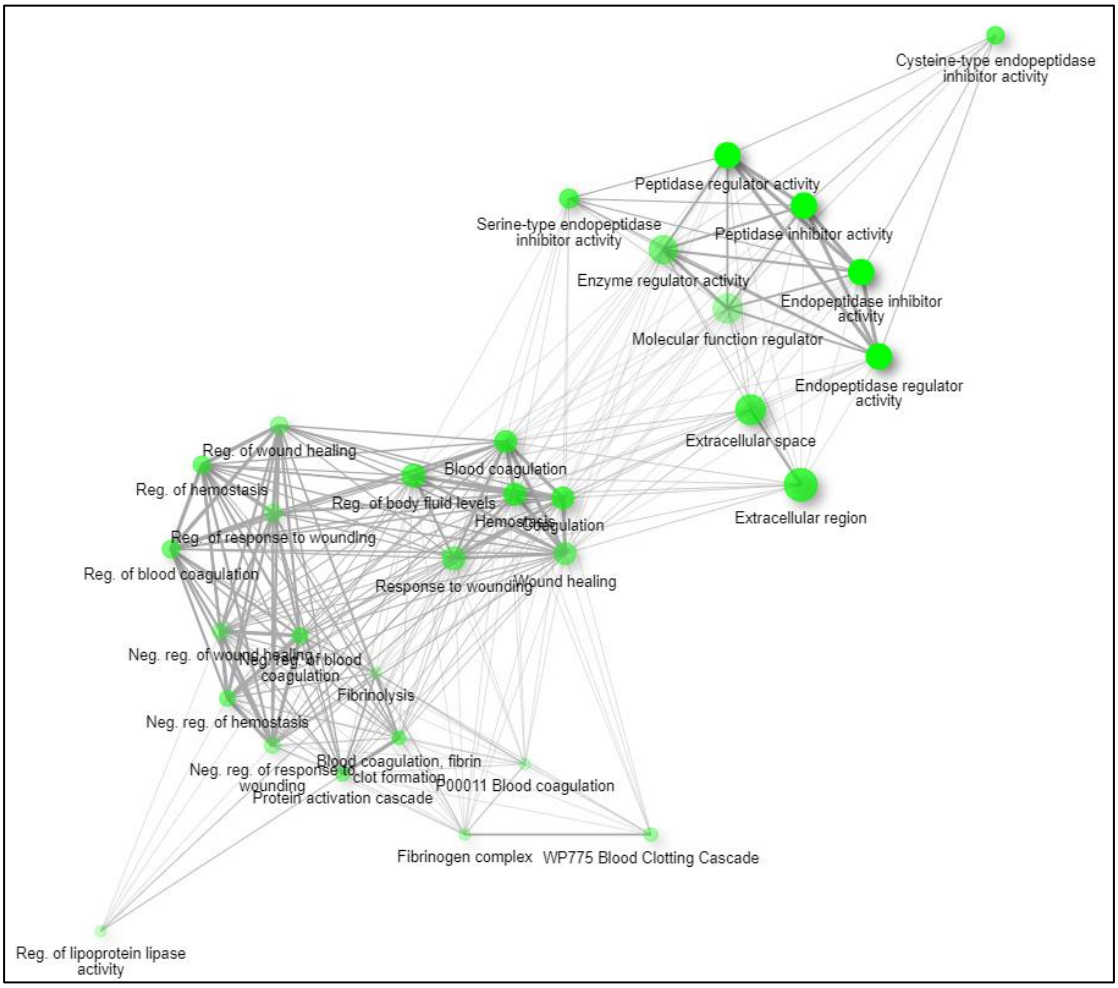


Fig. 4.23 Pathway-network of up and down-regulated genes in right ovary

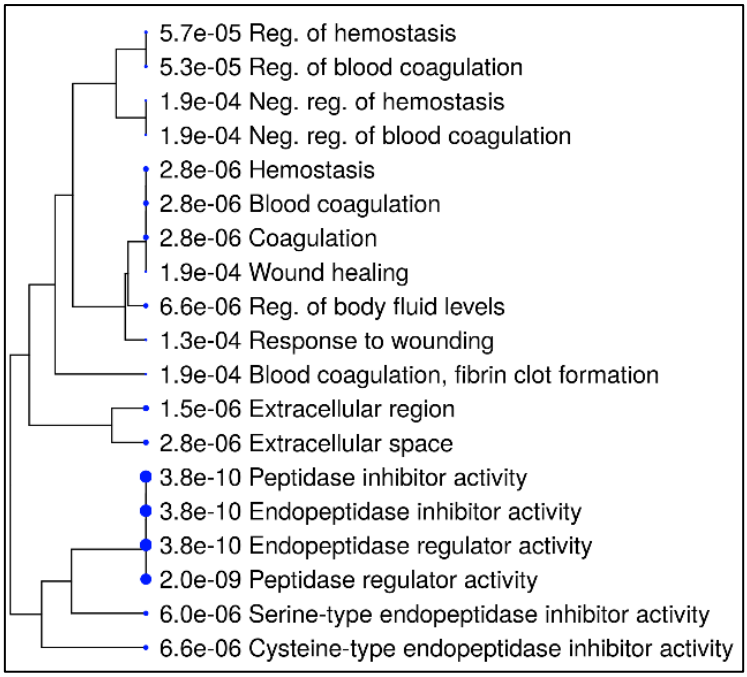


Fig. 4.24 Pathway tree-plot of up and down-regulated genes in right ovary

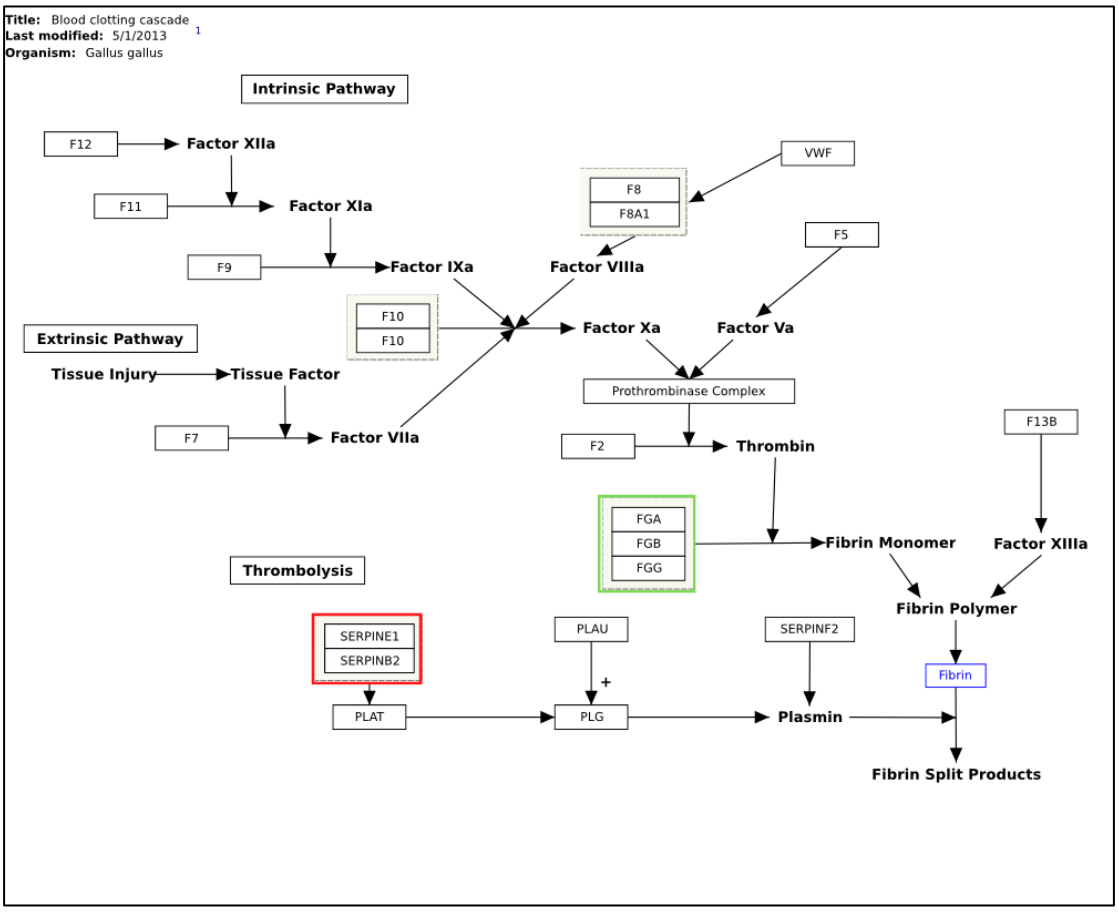


Fig. 4.25 Blood Clotting Cascade pathway

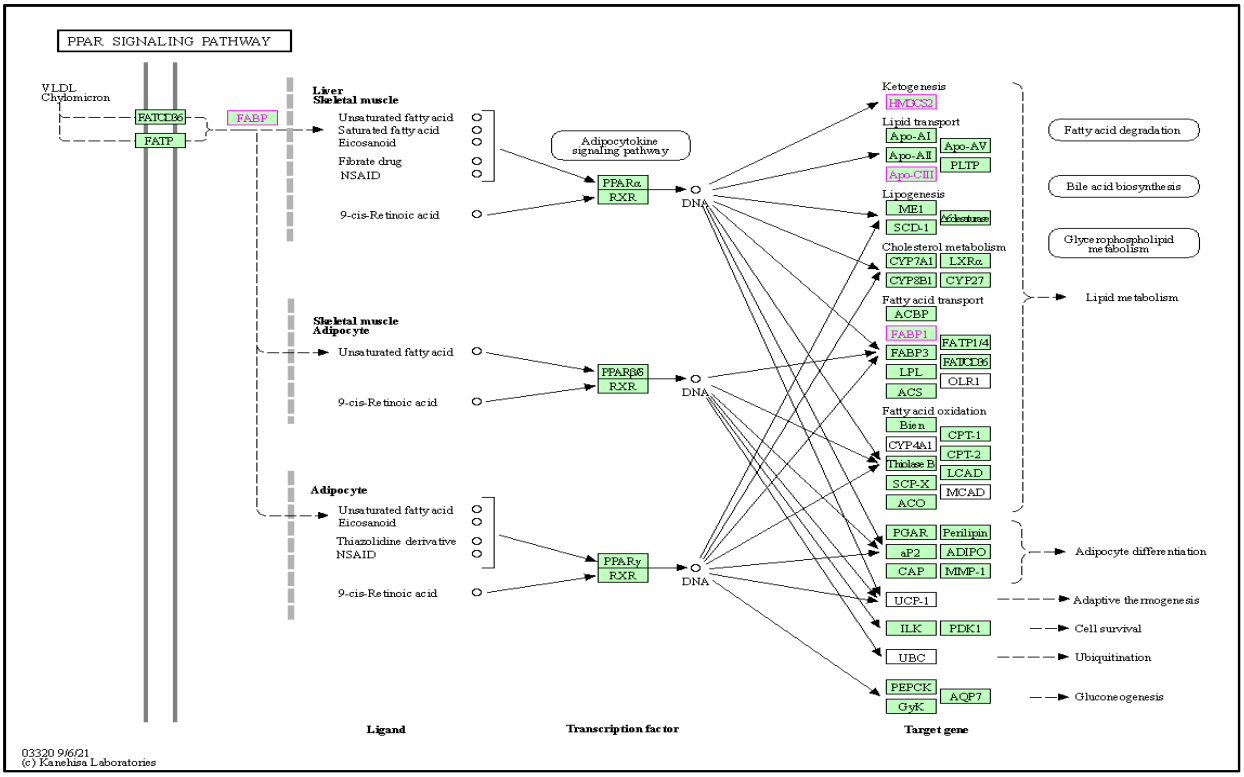


Fig. 4.26 PPAR signalling pathway

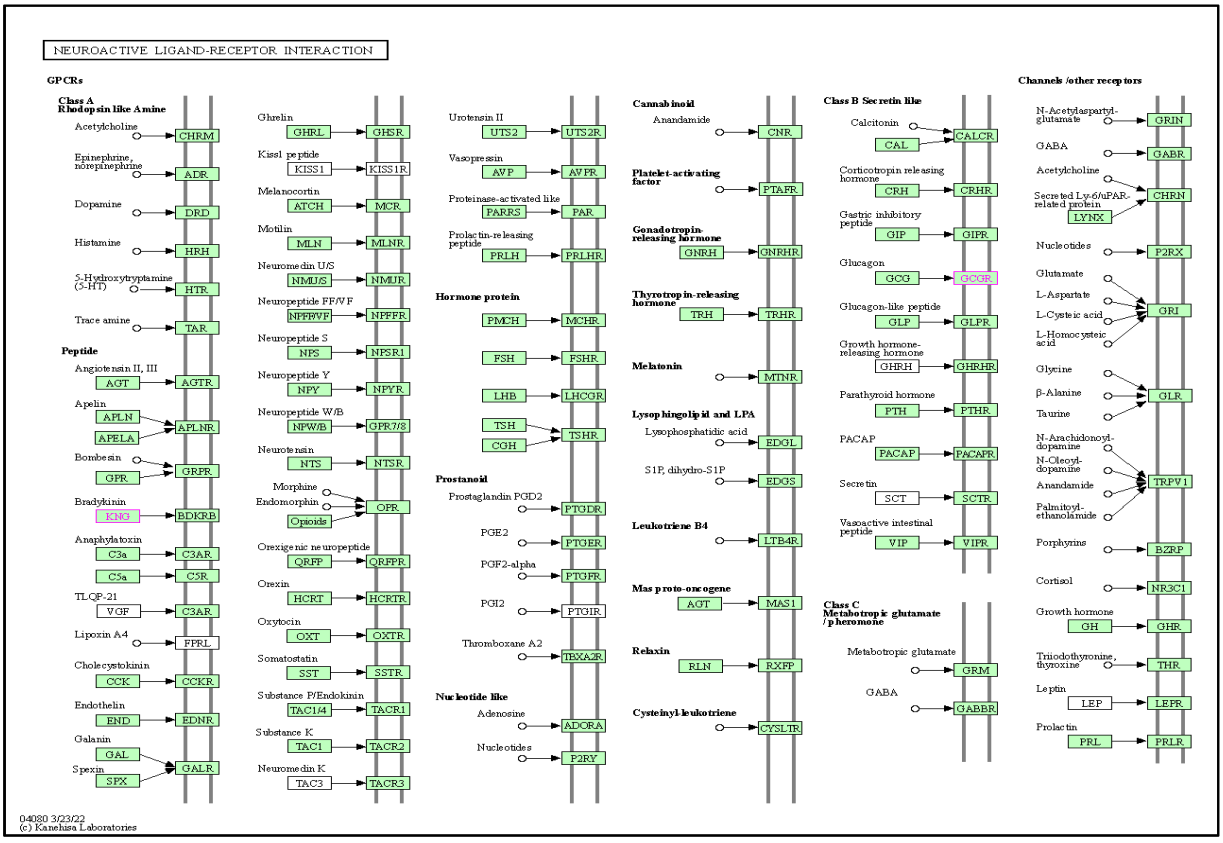


Fig. 4.27 Neuroactive ligand-receptor interaction

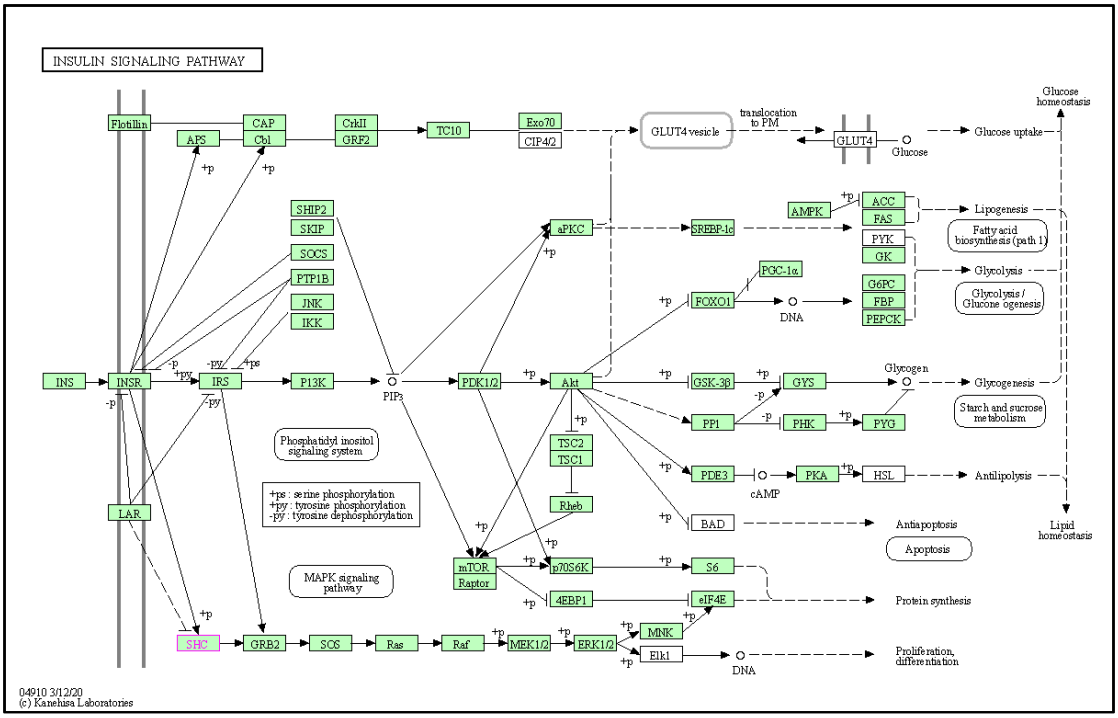


Fig. 4.28 Insulin signalling pathway (KEGG:04910)

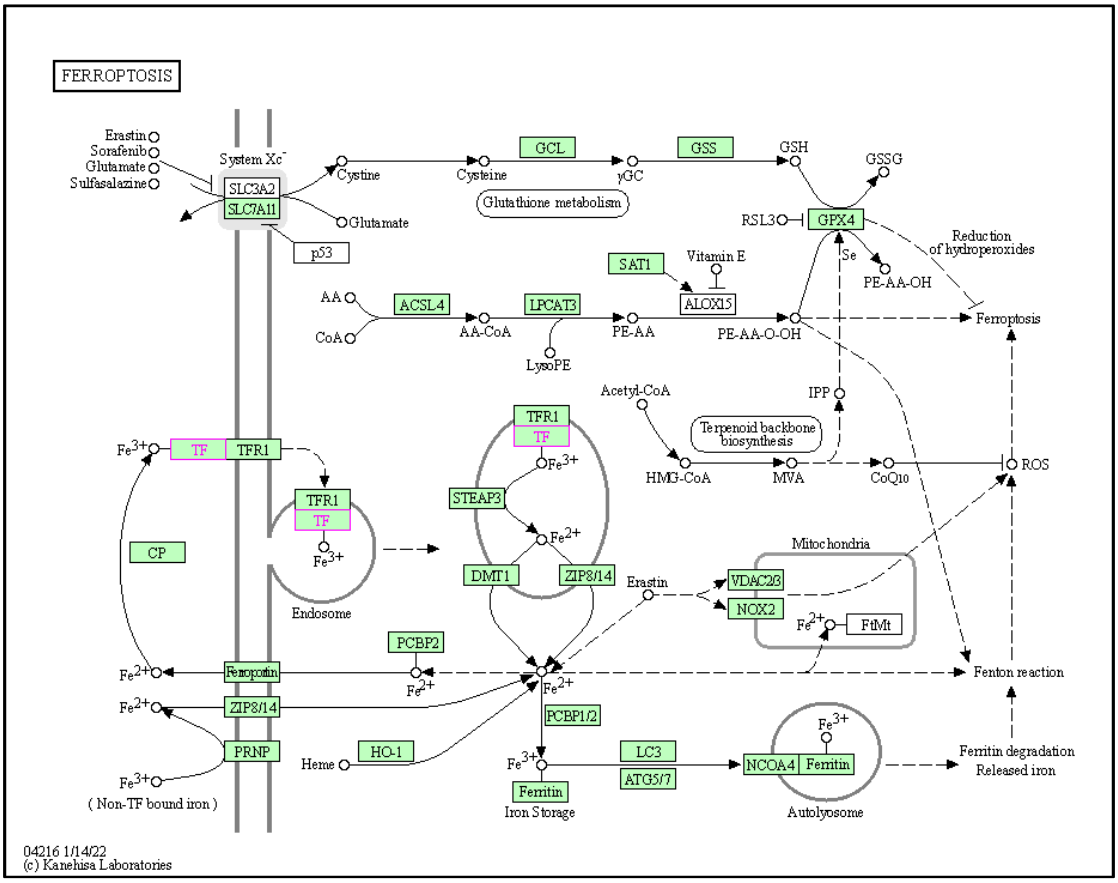


Fig. 4.29 Ferroptosis (KEGG:04216)

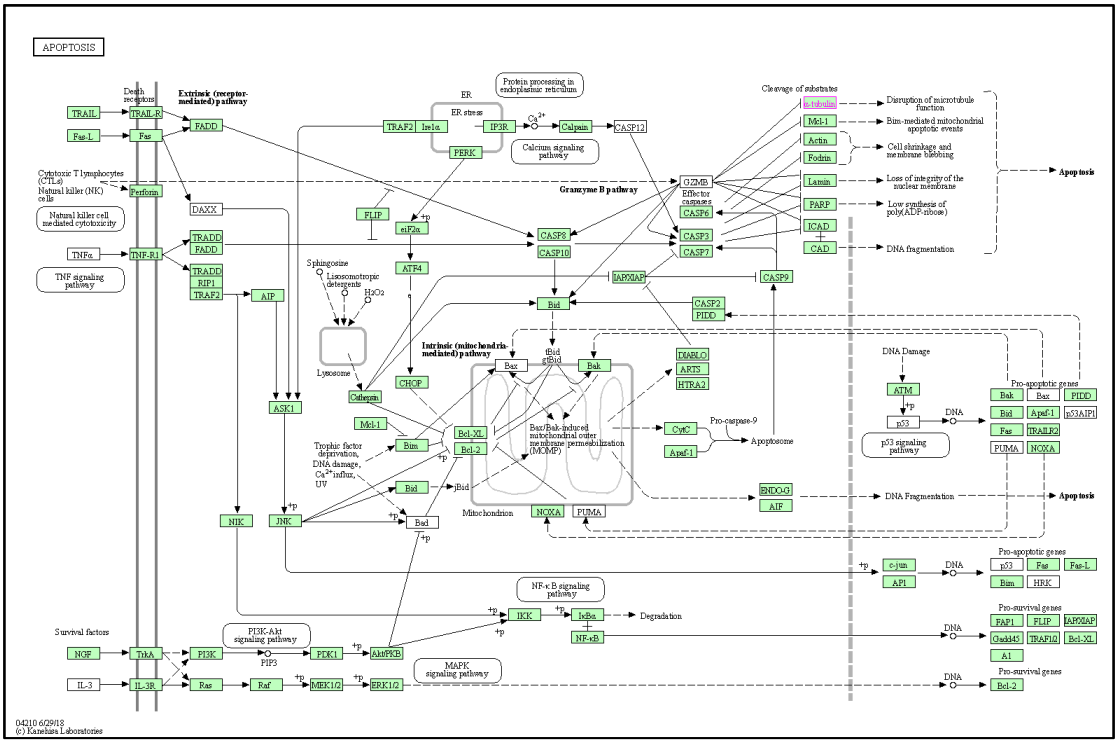


Fig. 4.30 LOC768589, TUBA3E involved in Apoptosis (KEGG:04210)

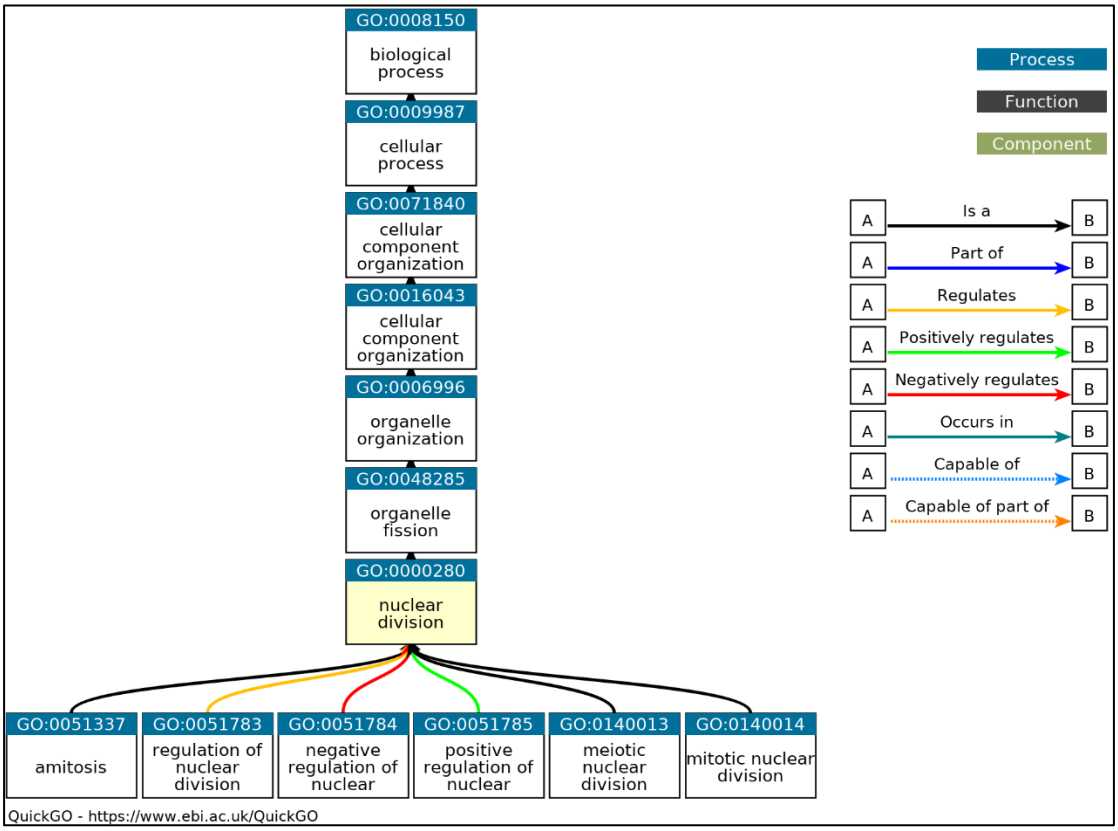


Fig. 4.31 Nuclear division (GO:0000280)

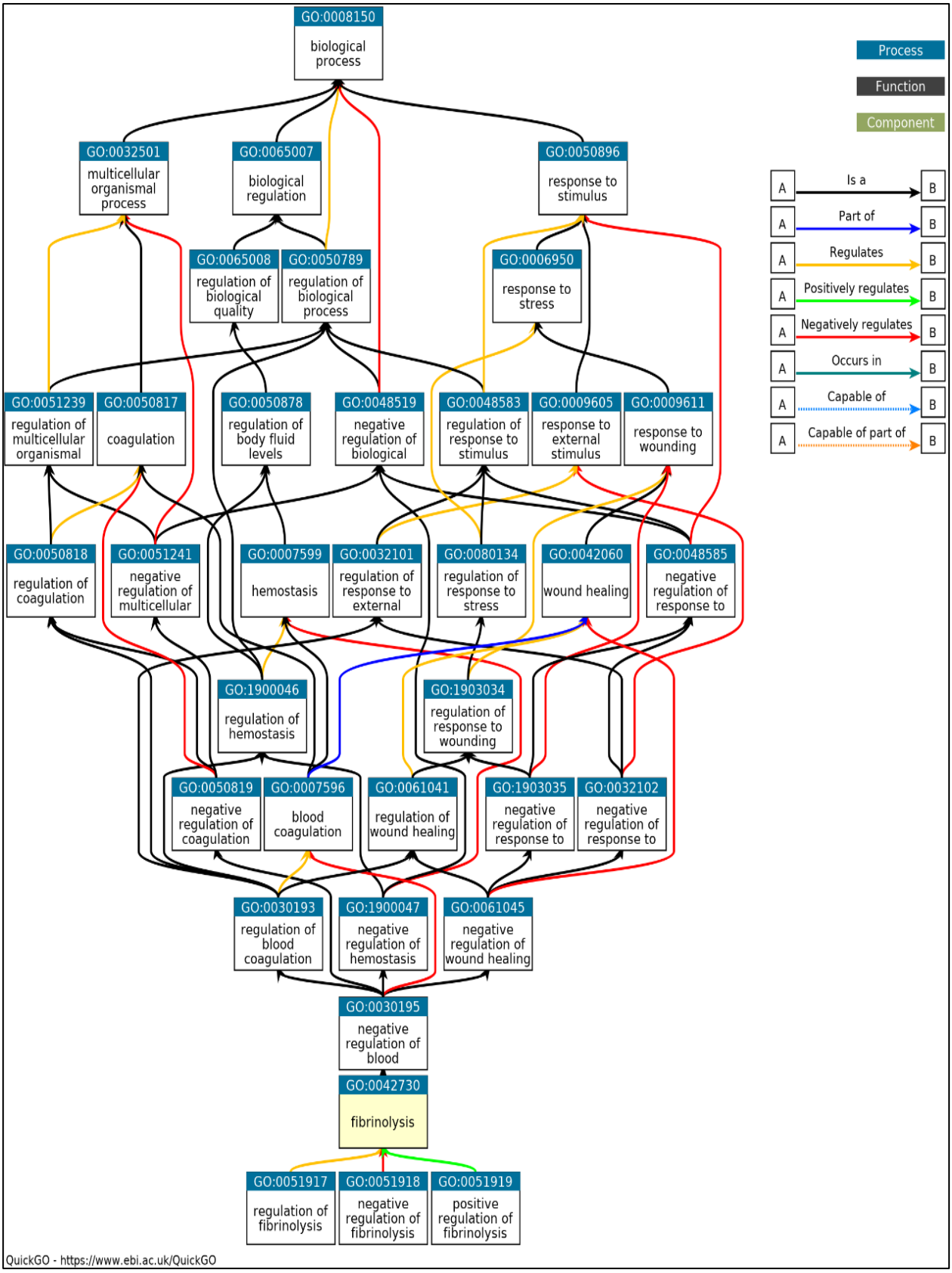


Fig. 4.32 Fibrinolysis (GO:0042730)

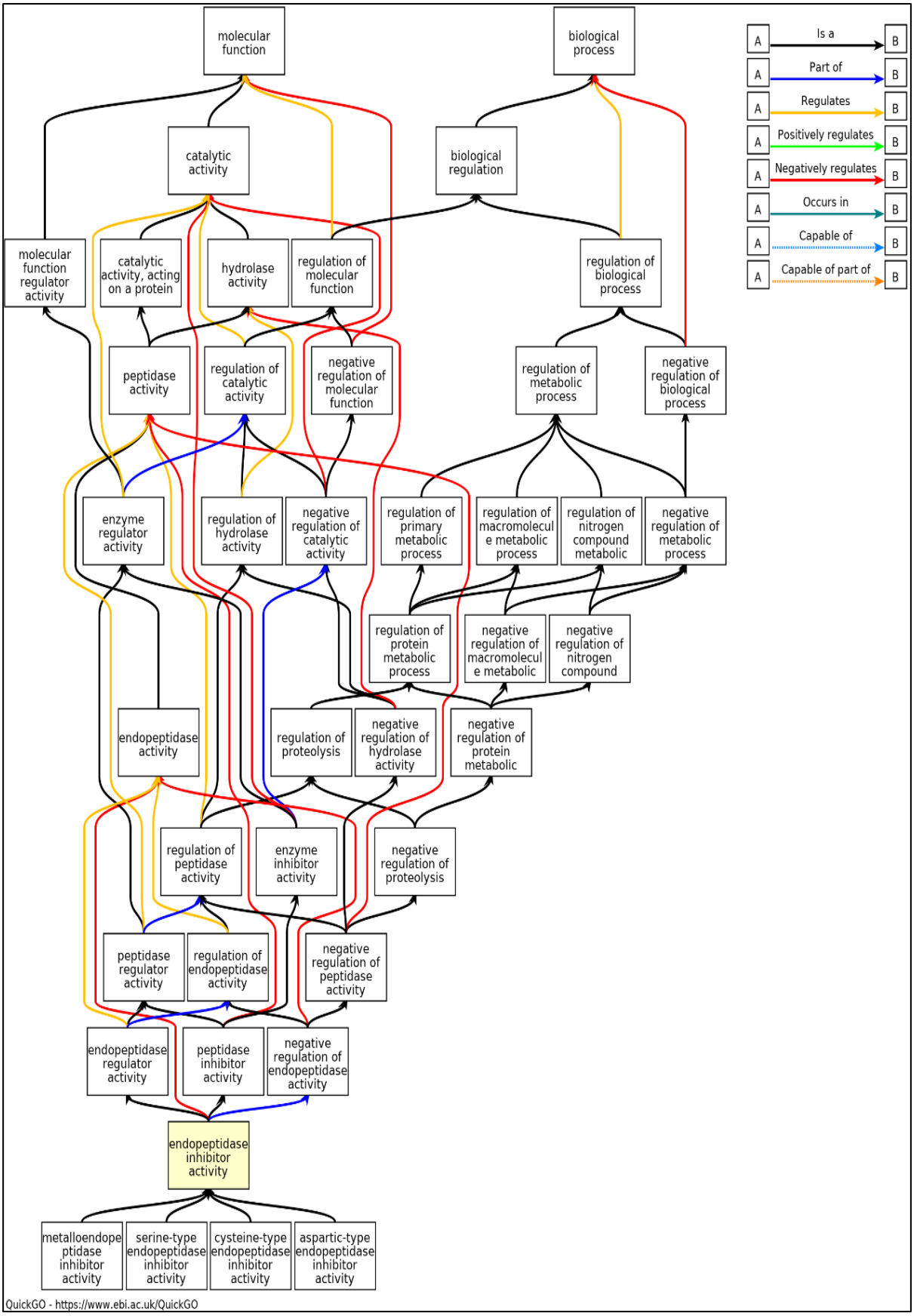


Fig. 4.33 Endopeptidase inhibitory activity (GO:0004866)

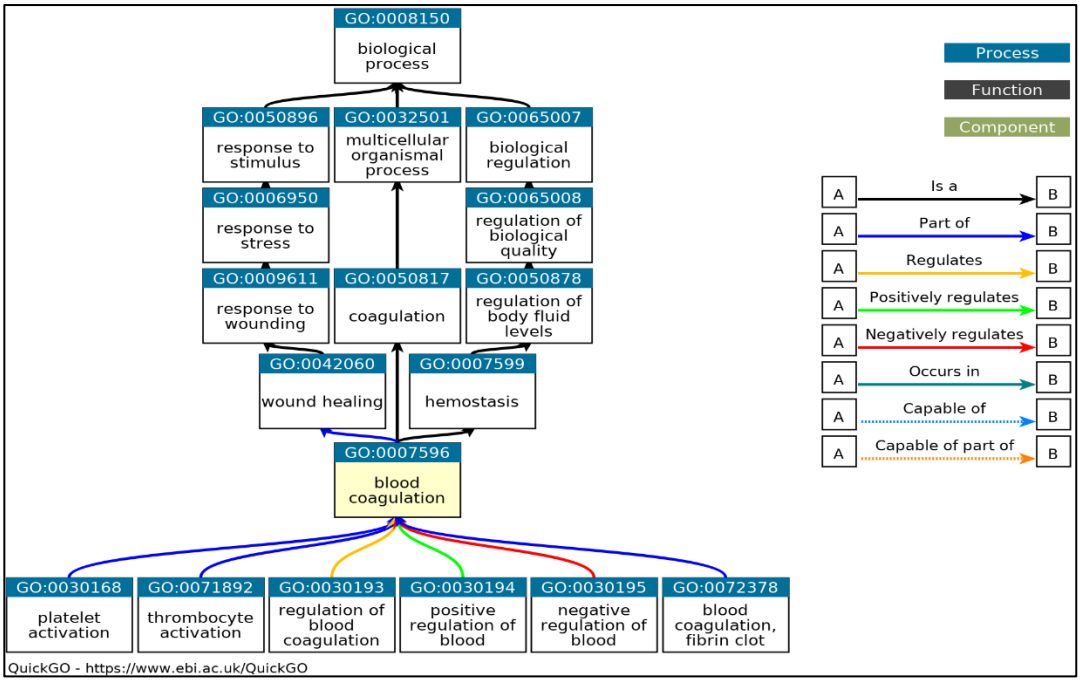


Fig. 4.34 Blood coagulation (GO:007596)

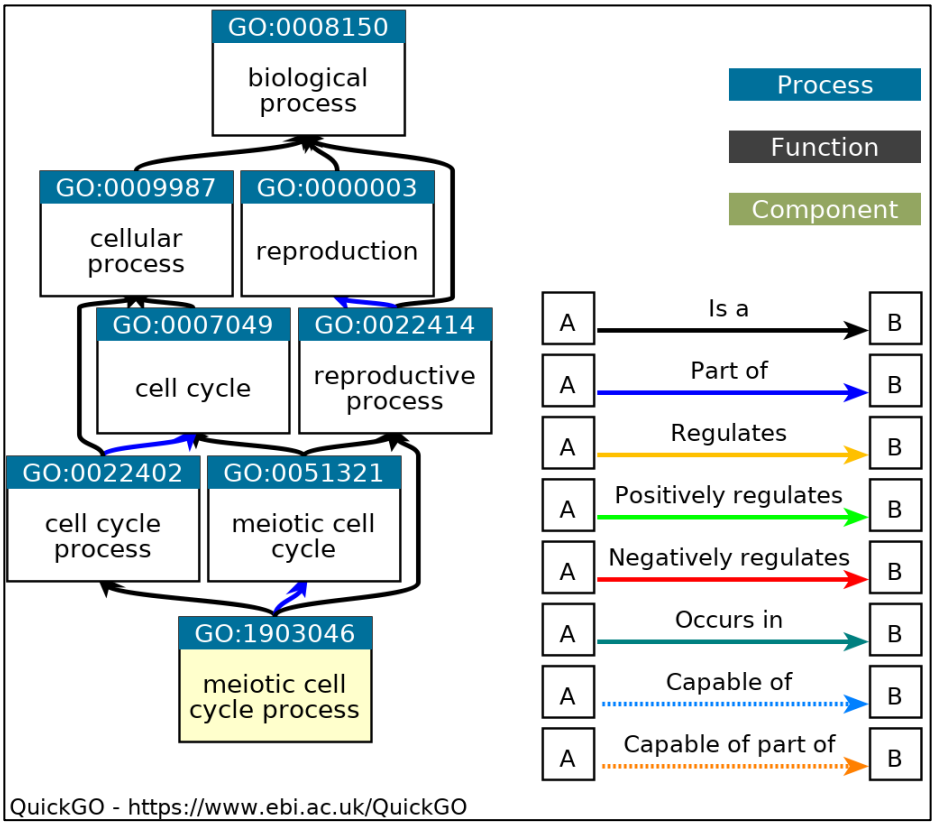


Fig. 4.35 Meiotic cell cycle process (GO:1903046)

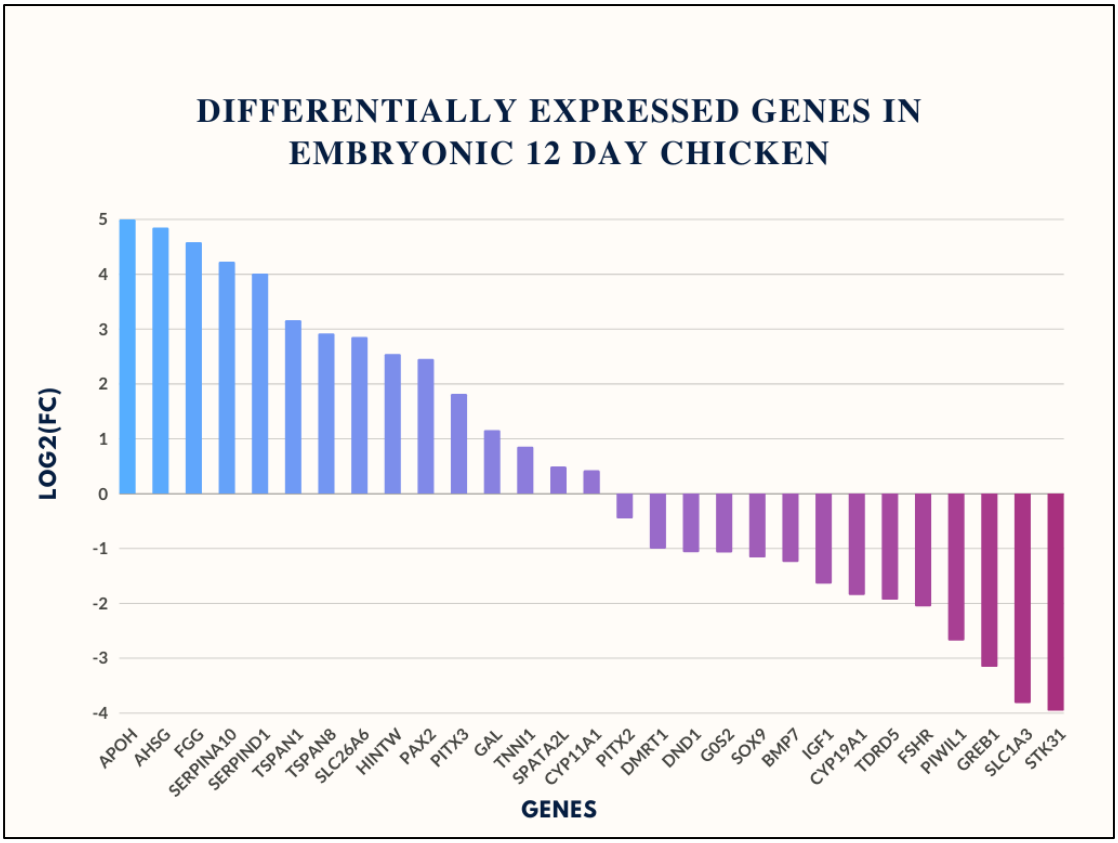


Fig. 4.36 Reported genes and their fold-changes in Transcriptome analysis

Table 4.11 ShinyGO pathway enrichment analysis for up and down-regulated genes

Sl No	Pathway	Genes
1	Blood coagulation, fibrin clot formation	SERPINC1 FGG APOH
2	Cysteine-type endopeptidase inhibitor activity	TF AHSG FETUB HRG KNG1
3	Neg. reg. of blood coagulation	PROC FGG APOH KNG1
4	Neg. reg. of hemostasis	PROC FGG APOH KNG1
5	Reg. of blood coagulation	PROC SERPINC1 FGG APOH KNG1
6	Reg. of hemostasis	PROC SERPINC1 FGG APOH KNG1
7	Serine-type endopeptidase inhibitor activity	SERPINC1 ITIH2 OVALY SERPINA10 AMBP SERPIND1
8	Endopeptidase inhibitor activity	SERPINC1 TF ITIH2 AHSG FETUB HRG OVALY SERPINA10 AMBP SERPIND1 KNG1
9	Peptidase inhibitor activity	SERPINC1 TF ITIH2 AHSG FETUB HRG OVALY SERPINA10 AMBP SERPIND1 KNG1
10	Endopeptidase regulator activity	SERPINC1 TF ITIH2 AHSG FETUB HRG OVALY SERPINA10 AMBP SERPIND1 KNG1
11	Peptidase regulator activity	SERPINC1 TF ITIH2 AHSG FETUB HRG OVALY SERPINA10 AMBP SERPIND1 KNG1
12	Blood coagulation	PROC SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1
13	Hemostasis	PROC SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1

Sl No	Pathway	Genes
14	Coagulation	PROC SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1
15	Reg. of body fluid levels	PROC AQP3 SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1
16	Wound healing	PROC SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1
17	Response to wounding	PROC SLC1A3 SERPINC1 FGA FGG SERPINA10 APOH SERPIND1 KNG1
18	Extracellular space	SERPINC1 TF AHSG FGA FGG SPIA4 ANGPTL3 GC ZPLD1 OVALY SERPINB2 SERPINA10 APOH ALB SERPIND1 KNG1
19	Extracellular region	PROC SPP2 SERPINC1 TF AHSG FGA FGG SPIA4 ANGPTL3 GC ZPLD1 OVALY SERPINB2 SERPINA10 RNP APOC3 APOH ALB SERPIND1 KNG1

Table 4.12 g:Profiler pathway enrichment analysis for up and down regulated genes

Sl No	Source	Term Name	Term id	Intersections
1	GO:MF	Endopeptidase inhibitor activity	GO:0004866	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,SERPINB2,OVALY
2	GO:MF	Peptidase inhibitor activity	GO:0030414	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,SERPINB2,OVALY

3	GO:MF	Endopeptidase regulator activity	GO:0061135	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,SERPINB2,OVALY
4	GO:MF	Peptidase regulator activity	GO:0061134	FETUB,AHSG,SPIA4,SERPINC1,HRG,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,SERPINB2,OVALY
5	GO:MF	Enzyme inhibitor activity	GO:0004857	APOV1,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,SERPINB2,OVALY
6	GO:MF	Serine-type endopeptidase inhibitor activity	GO:0004867	SPIA4,SERPINC1,ITIH2,SERPINA10,AMBP,SERPIND1,SERPINB2,OVALY
7	GO:MF	Enzyme regulator activity	GO:0030234	APOV1,APOH,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,RGS7,SERPINB2,TOM1L1,OVALY
8	GO:MF	Cysteine-type endopeptidase inhibitor activity	GO:0004869	FETUB,AHSG,HRG,KNG1
9	GO:MF	Molecular function regulator activity	GO:0098772	APOV1,APOH,FETUB,AHSG,SPIA4,SERPINC1,HRG,ANGPTL3,ITIH2,SERPINA10,AMBP,SERPIND1,KNG1,RNP,RGS7,SERPINB2,TOM1L1,OVALY
10	GO:BP	Blood coagulation	GO:0007596	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
11	GO:BP	Hemostasis	GO:0007599	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
12	GO:BP	Coagulation	GO:0050817	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
13	GO:BP	Regulation of body fluid levels	GO:0050878	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,AQP3,SERPIND1,PROC,KNG1
14	GO:BP	Regulation of blood coagulation	GO:0030193	CPB2,APOH,SERPINC1,FGG,PROC,KNG1

15	GO:BP	Regulation of hemostasis	GO:1900046	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
16	GO:BP	Regulation of coagulation	GO:0050818	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
17	GO:BP	Negative regulation of hemostasis	GO:1900047	CPB2,APOH,FGG,PROC,KNG1
18	GO:BP	Negative regulation of blood coagulation	GO:0030195	CPB2,APOH,FGG,PROC,KNG1
19	GO:BP	Negative regulation of coagulation	GO:0050819	CPB2,APOH,FGG,PROC,KNG1
20	GO:BP	Response to wounding	GO:0009611	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1,SLC1A3
21	GO:BP	Wound healing	GO:0042060	CPB2,APOH,SERPINC1,FGA,FGG,SERPINA10,SERPIND1,PROC,KNG1
22	GO:BP	Negative regulation of wound healing	GO:0061045	CPB2,APOH,FGG,PROC,KNG1
23	GO:BP	Regulation of wound healing	GO:0061041	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
24	GO:BP	Negative regulation of response to wounding	GO:1903035	CPB2,APOH,FGG,PROC,KNG1
25	GO:BP	Regulation of response to wounding	GO:1903034	CPB2,APOH,SERPINC1,FGG,PROC,KNG1
26	GO:BP	Fibrinolysis	GO:0042730	CPB2,APOH,FGG
27	GO:BP	Blood coagulation, fibrin clot formation	GO:0072378	APOH,SERPINC1,FGG
28	GO:BP	Pirna metabolic process	GO:0034587	TDRD9,TDRD1,ASZ1

29	GO:BP	DNA methylation involved in gamete generation	GO:0043046	TDRD9,TDRD1,ASZ1
30	GO:BP	Protein activation cascade	GO:0072376	APOH,SERPINC1,FGG
31	GO:CC	Extracellular region	GO:0005576	ORM1,APOV1,SPP2,RBP4A,APOH,AHSG,SPIA4,GC,SERPINC1,FGA,TF,TMPRSS6,FGG,ANGPTL3,SERPINA10,ALB,APOC3,SERPIND1,PROC,KNG1,MBL2,RNP,SERPINB2,ZPLD1,OVALY
32	GO:CC	Extracellular space	GO:0005615	ORM1,APOV1,RBP4A,APOH,AHSG,SPIA4,GC,SERPINC1,FGA,TF,TMPRSS6,FGG,ANGPTL3,SERPINA10,ALB,SERPIND1,KNG1,MBL2,SERPINB2,ZPLD1,OVALY
33	GO:CC	P granule	GO:0043186	TDRD9,TDRD1,ASZ1
34	GO:CC	Pole plasm	GO:0045495	TDRD9,TDRD1,ASZ1
35	GO:CC	Germ plasm	GO:0060293	TDRD9,TDRD1,ASZ1
36	GO:CC	Pi-body	GO:0071546	TDRD1,ASZ1
37	GO:CC	Chylomicron	GO:0042627	APOV1,APOH
38	REAC	Post-translational protein phosphorylation	REAC:R-GGA-8957275	SPP2,AHSG,TF,FGG,ITIH2,SERPINA10,SERPIND1,PROC,KNG1
39	REAC	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (igfbps)	REAC:R-GGA-381426	SPP2,AHSG,TF,FGG,ITIH2,SERPINA10,SERPIND1,PROC,KNG1

40	REAC	Platelet degranulation	REAC:R-GGA-114608	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1
41	REAC	Response to elevated platelet cytosolic Ca2+	REAC:R-GGA-76005	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1
42	REAC	Intrinsic Pathway of Fibrin Clot Formation	REAC:R-GGA-140837	SPIA4,SERPIND1,PROC,KNG1
43	REAC	Formation of Fibrin Clot (Clotting Cascade)	REAC:R-GGA-140877	SPIA4,SERPIND1,PROC,KNG1
44	REAC	Platelet activation, signaling and aggregation	REAC:R-GGA-76002	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,KNG1
45	REAC	Common Pathway of Fibrin Clot Formation	REAC:R-GGA-140875	SPIA4,SERPIND1,PROC
46	REAC	Hemostasis	REAC:R-GGA-109582	ORM1,SPP2,APOH,AHSG,SPIA4,TF,HRG,SERPIND1,PROC,KNG1
47	WP	Blood clotting cascade	WP:WP775	FGA,FGG,SERPINB2
48	HP	Non-obstructive azoospermia	HP:0011961	TDRD9,DAZL,SYCP3,TERB1
49	HP	Recurrent thromboembolism	HP:0004831	SERPINC1,HRG
Notes: BP - Biological Process, MF - Molecular Function, CC - Cellular Component, KEGG - Kyoto Encyclopedia of Genes and Genomes, REAC - Reactome, WP – WikiPathways, HP - human phenotype ontology				

4.2 SEXING OF CHICKEN EMBRYOS

Identification of male and female embryonic chickens was identified using SWIM primers. Identification of male and female chicken embryos depending upon the morphological characteristics is difficult at the embryonic 12th day stage, as the gonads are not completely differentiated into testis and ovaries. A total of 20 embryonic chickens were selected for the experiment. Among them, 12 were verified as females and 8 were verified as males. The female gonads were processed for future evaluation of gene expression analysis and male gonads were discarded.

4.2.1 Isolation of genomic DNA and determination of its quantity and quality

The embryonic liver samples were collected aseptically and stored at -80^{o}C. The good-quality genomic DNA was isolated through the phenol-chloroform extraction method. The Purity and concentration of isolated DNA were evaluated based on OD values (260:280) through Nano drop. The quality of the DNA was checked for any shearing by agarose gel electrophoresis (0.8%) and appeared as a single band. The DNA was diluted with nuclease-free water to adjust the concentration between 50-100 ng/µl and used as a template in the PCR reaction.

4.2.3 PCR amplification by SWIM primers

The PCR reaction was set with sex-specific primer - SWIM for the amplification in 0.2 ml capacity flat PCR tubes, using a thermal cycler (Himedia). 10µl of the PCR reactions were analysed by gel electrophoresis system and visualised under Gel Documentation (Figure 4.37)

4.3 VALIDATION OF TRANSCRIPTOME ANALYSIS

Transcriptome analysis of the right and left ovary of chicken during the embryonic 12th day revealed that the total number of significantly up and down-regulated differentially expressed genes between embryonic left and right chicken ovaries was found to be 9152. Similarly, the total number of significantly up-regulated genes by a fold change above +2 in the right ovary during E12 was 479 and the significantly down-regulated genes by fold change below -2 in the right ovary during E12 were 227.

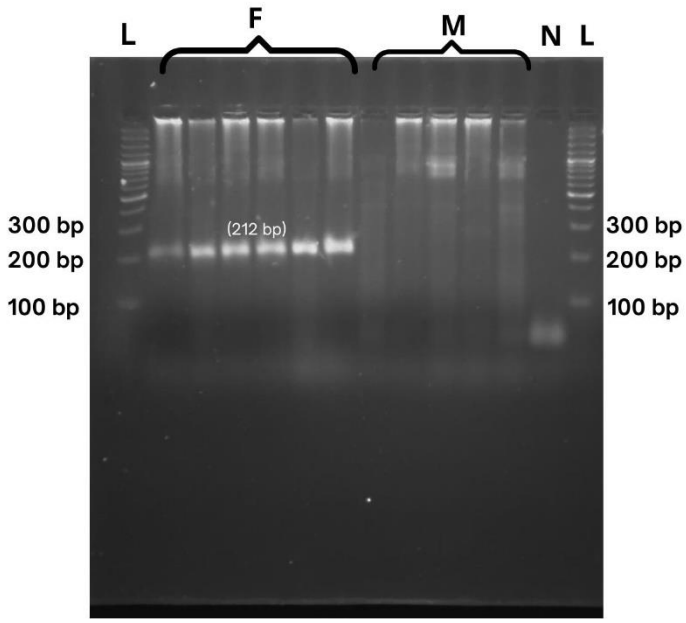


Fig. 4.37 Agarose gel electrophoresis showing PCR amplified product of SWIM (212 bp). L- Ladder, F- Female, M-Male, N- Neutral.

Validation was performed with the two most influenced up-regulated and two most impacted down-regulated genes to support the above-mentioned transcriptome analysis. FETUB (Fetuin-B) was chosen for validation of up-regulated genes because it is involved in gene ontology molecular functions like serine-type endopeptidase inhibitor activity and potent inhibitor of ovastacin (Dietzel E *et al.*, 2016). LEAP2 (Liver Enriched Antimicrobial Peptide 2), which is known as an endogenous blocker of growth hormone secretagogue receptor 1a (GHS-R1a) (Lu *et al.*, 2021). A down-regulated gene called DDX4 (DEAD-Box Helicase 4), an RNA helicase and homolog of VASA proteins, is necessary for the correct growth of primordial germ cells (Aduma *et al.*, 2019). The RNA-binding protein DAZL (deleted in azoospermia-like) is a crucial regulator of germ cell development and entry into meiosis (Chen *et al.*, 2021).

4.3.1 Isolation of total RNA from tissues

A total of 20 fertile eggs of Kadaknath birds on the 12th day of incubation were randomly chosen at ICAR - Directorate of Poultry Research, Rajendranagar, Hyderabad. Embryos were sacrificed at the National Bureau of Plant Genetic Resources Regional Station (NBPGR), Rajendranager for gonadal and liver tissue collection. The right regressing ovary cannot be collected with naked eyes due to its smaller size, complete collection was done with the help of a stereo microscope (Leica) which is

available at NBPGR, Rajendranagar. The gonadal tissues were stored in RNA later and the liver was stored in PBS at -80^{o}C. Total RNA was isolated using TRIzol reagent from the right and left ovary samples.

4.3.2 Concentration, Quality and Quantity of RNA

The concentration, quality and quantity of RNA were assessed using the NanoDrop spectrophotometer. The concentration of the RNA samples was measured in ng/µl using one microlitre of the isolated RNA. Spectrophotometric readings at OD260/OD280 represented the purity of the total RNA and ranged between 1.9 and 2.2 indicating purity and lack of DNA and protein contamination. The concentration of RNA, OD260/280 and OD 260/230 were mentioned in Table 4.13.

Table 4.13: Details of concentration of RNA obtained, OD260/280 and OD 260/230 ratios

Sl. No	Tissue Sample	RNA Sample	The concentration of RNA (ng/µl)	260/280	260/230
1	Left Ovary 1	L1	1487.8	1.91	2.18
2	Left Ovary 2				
3	Left Ovary 3				
4	Left Ovary 4	L2	1384.2	1.89	2.12
5	Left Ovary 5				
6	Left Ovary 6				
7	Left Ovary 7	L3	1570.4	1.93	2.02
8	Left Ovary 8				
9	Left Ovary 9				
10	Left Ovary 10	L4	1122.3	1.88	2.09
11	Left Ovary 11				
12	Left Ovary 12				
13	Right Ovary 1	R1	968.5	1.89	2.16
14	Right Ovary 2				
15	Right Ovary 3				
16	Right Ovary 4	R2	1078.4	1.94	2.17

17	Right Ovary 5				
18	Right Ovary 6				
19	Right Ovary 7	R3	1210.7	1.93	2.06
20	Right Ovary 8				
21	Right Ovary 9				
22	Right Ovary 10	R4	1198.4	1.88	2.09
23	Right Ovary 11				
24	Right Ovary 12				

4.3.3 Synthesis of cDNA

The first-strand cDNA was synthesized from total RNA (1µg) samples using oligo primers supplied with Puregene First Strand cDNA Synthesis Kit (Genetix Biotech) which follows a method of reverse transcription. Due to the frailty of cDNA, it was used immediately for further downstream processing.

4.3.4 Real-Time Quantitative PCR (RT-qPCR)

The RT-qPCR conditions for target genes (FETUB, LEAP2, DAZL and DDX4) and reference gene (β actin) were standardized using different combinations of dNTPs, primers and annealing temperatures. The reaction mixture (Table 3.10) and thermal cycling conditions (Table 3.11) of RT-PCR were found to be optimal. Duplicates of each sample were kept for analysis. RT-qPCR was performed to assess the differential expression of FETUB, LEAP2, DAZL and DDX4 in both left and right ovarian tissues using RT-qPCR SYBR Green assay using specific primers (Table 3.8 and Table 3.9) with cDNA as template. The amplification plot and melting curves generated during RT-qPCR for β actin, FETUB, LEAP2, DAZL and DDX4 in both left and right ovaries were shown in Fig. 4.38a, 4.38b, 4.39a, 4.39b, 4.40a, 4.40b, 4.41a, 4.41b, 4.42a and 4.42b.

4.3.5 Expression of FETUB gene

The expression of the FETUB gene was studied in both left and right ovarian tissues of Kadaknath chicken embryos on the 12th day of incubation. Among the tissues, the highest expression of FETUB was noticed in the right ovaries on the 12th day of

incubation while the least expression was noticed in the left ovaries of the same age group.

4.3.5.1 Relative expression of FETUB gene in left and right ovaries

The higher expression of the FETUB gene was seen in the right ovaries when compared to the left ovaries. The mean C_T value for FETUB and β actin, ΔC_T and fold change ($2^{-\Delta\Delta CT}$) are presented in Table 4.14. The relative tissue expressions of FETUB mRNA were normalized with the reference gene (β actin). The gene expression in the left ovaries was taken as a control. The expression was higher in the right ovaries by 5.33 folds compared to the left ovaries. [No. of samples (N) =4 left ovaries and 4 right ovaries]

Table 4.14: Relative expression profile of FETUB gene in left and right ovaries

Sl. No	Tissue	Mean C_T ± SE		ΔC_T ± SE	Fold change ($2^{-\Delta\Delta CT}$)
		FETUB	β actin		
1.	Left ovaries (Control)	24.57±1.95	20.70±1.68	3.87±1.23	1
2.	Right ovaries	22.61±0.73	20.30±0.78	2.32±1.25	5.33

4.3.6 Expression of LEAP2 gene

The expression of the LEAP2 gene was studied in both left and right ovarian tissues of Kadaknath chicken embryos on the 12th day of incubation. Among the tissues, the highest expression of LEAP2 was noticed in the right ovaries on the 12th day of incubation while the least expression was noticed in left ovaries of the same age group.

4.3.6.1 Relative expression of LEAP2 gene in left and right ovaries

The higher expression of the LEAP2 gene was seen in the right ovaries when compared to the left ovaries. The mean C_T value for LEAP2 and β actin, ΔC_T and fold change ($2^{-\Delta\Delta CT}$) are presented in Table 4.15. The relative tissue expressions of LEAP2 mRNA were normalized with the reference gene (β actin). The gene expression in the left ovaries was taken as a control. The expression was higher in the right ovaries by

2.27 folds compared to the left ovaries. [No. of samples (N) =4 left ovaries and 4 right ovaries]

Table 4.15: Relative expression profile of LEAP2 gene in left and right ovaries

Sl. No	Tissue	Mean C_T ± SE		ΔC_T ± SE	Fold change ($2^{-\Delta\Delta CT}$)
		LEAP2	β actin		
1.	Left ovaries (Control)	22.58±2.13	20.70±1.68	1.88±2.27	1
2.	Right ovaries	21.59±0.46	20.30±0.78	1.30±0.96	2.27

4.3.7 Expression of DAZL gene

The expression of the DAZL gene was studied in both left and right ovarian tissues of Kadaknath chicken embryos on the 12th day of incubation. Among the tissues, the lowest expression of DAZL was noticed in the right ovaries on the 12th day of incubation while the highest expression was noticed in the left ovaries of the same age group.

4.3.7.1 Relative expression of DAZL gene in left and right ovaries

The lowest expression of the DAZL gene was seen in the right ovaries when compared to the left ovaries. The mean C_T value for DAZL and β actin, ΔC_T and fold change ($2^{-\Delta\Delta CT}$) are presented in Table 4.16. The relative tissue expressions of DAZL mRNA were normalized with the reference gene (β actin). The gene expression in the left ovaries was taken as a control. The expression was lowest in the right ovaries by 0.15 folds compared to the left ovaries. [No. of samples (N) =4 left ovaries and 4 right ovaries]

Table 4.16: Relative expression profile of DAZL gene in left and right ovaries

Sl. No	Tissue	Mean C_T ± SE		ΔC_T ± SE	Fold change ($2^{-\Delta\Delta CT}$)
		DAZL	β actin		
1.	Left ovaries (Control)	18.53±1.36	20.30±1.68	-2.17±0.84	1

2.	Right ovaries	22.67±1.13	20.30±0.78	2.38±1.45	0.15

4.3.8 Expression of DDX4 gene

The expression of the DDX4 gene was studied in both left and right ovarian tissues of Kadaknath chicken embryos on the 12th day of incubation. Among the tissues, the lowest expression of DDX4 was noticed in the right ovaries on the 12th day of incubation while the highest expression was noticed in the left ovaries of the same age group.

4.3.7.1 Relative expression of DDX4 gene in left and right ovaries

The lowest expression of the DDX4 gene was seen in the right ovaries when compared to the left ovaries. The mean C_T value for DDX4 and β actin, ΔC_T and fold change ($2^{-\Delta\Delta CT}$) are presented in Table 4.17. The relative tissue expressions of DDX4 mRNA were normalized with the reference gene (β actin). The gene expression in the left ovaries was taken as a control. The expression was lowest in the right ovaries by 0.19 folds compared to the left ovaries. [No. of samples (N) =4 left ovaries and 4 right ovaries]

Table 4.17: Relative expression profile of DDX4 gene in left and right ovaries

Sl. No	Tissue	Mean C_T ± SE		ΔC_T ± SE	Fold change ($2^{-\Delta\Delta CT}$)
		DDX4	β actin		
1.	Left ovaries (Control)	19.433±1.73	20.30±1.68	-1.27±1.37	1
2.	Right ovaries	22.63±0.68	20.30±0.78	2.34±1.00	0.19

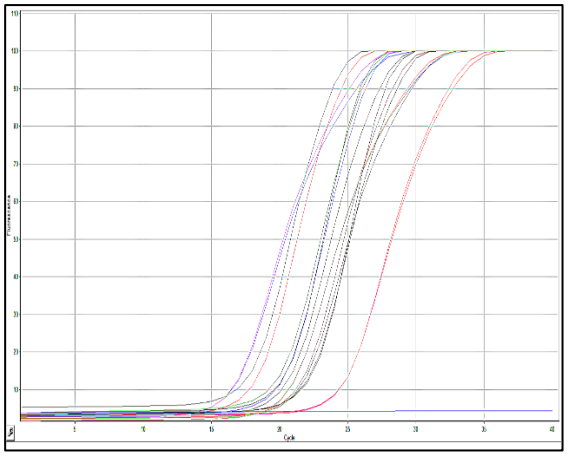


Fig. 4.38a Amplification plot of β actin

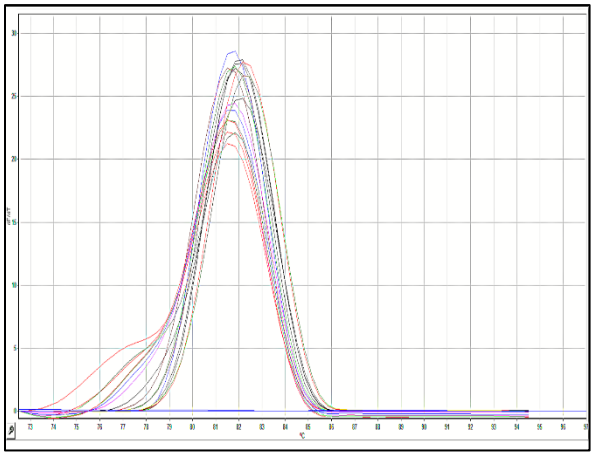


Fig. 4.38b Melting curve of β actin

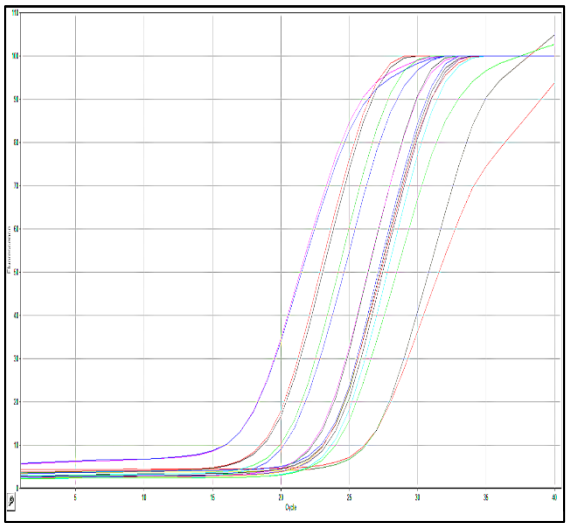


Fig. 4.39a Amplification plot of FETUB

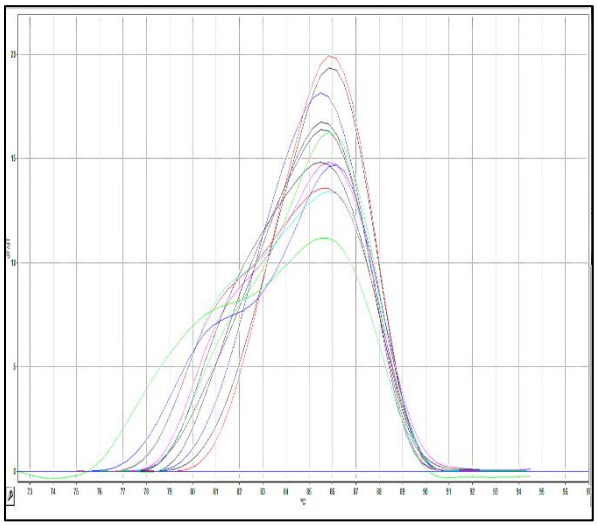


Fig. 4.39b Melting curve of FETUB

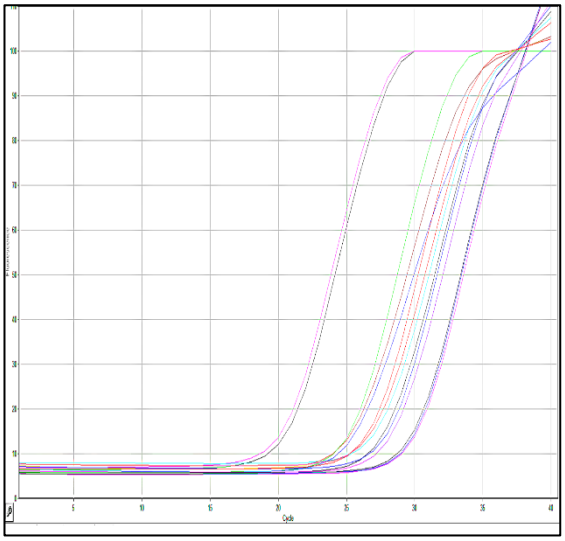


Fig. 4.40a Amplification plot of LEAP2

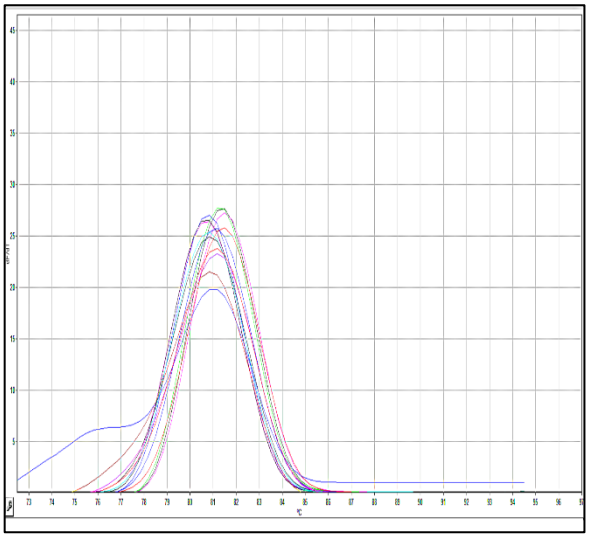


Fig. 4.40b Melting curve of LEAP2

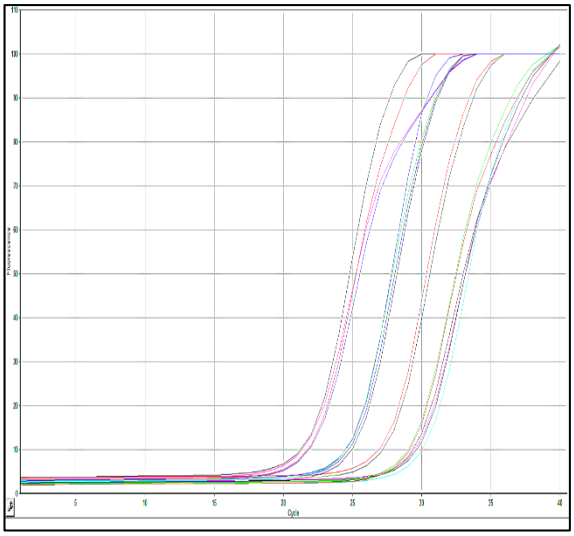


Fig. 4.41a Amplification plot of DAZL

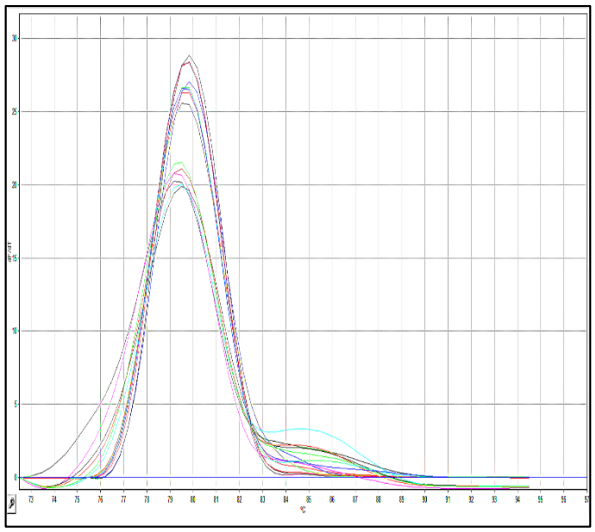


Fig. 4.41b Melting curve of DAZL

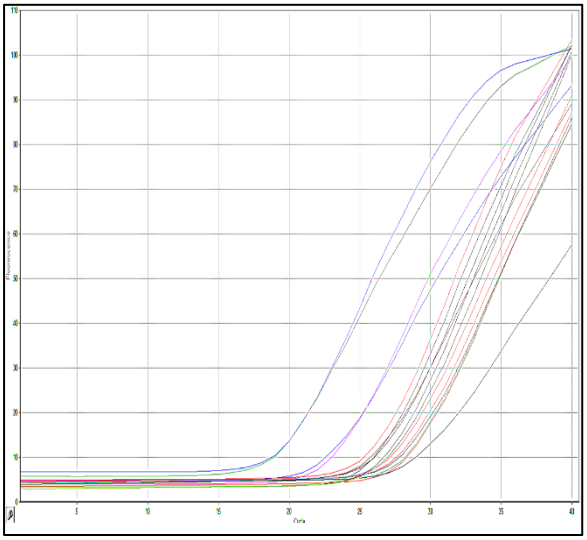


Fig. 4.42a Amplification plot of DDX4

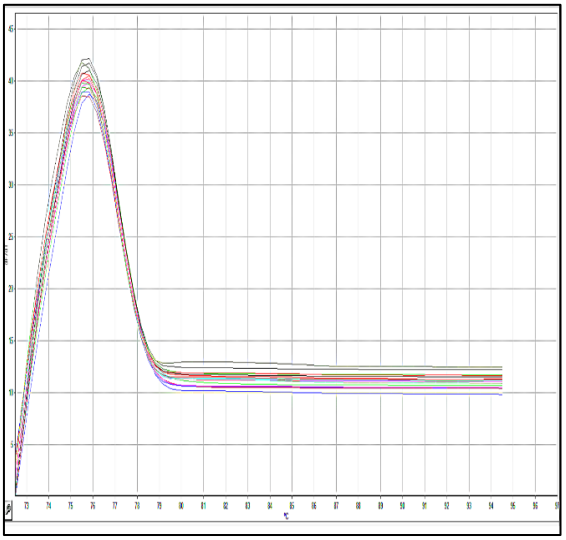


Fig. 4.42b Melting curve of DDX4

DISCUSSION

CHAPTER V

DISCUSSION

The chicken (*Gallus gallus domesticus*) is one of the most common domesticated poultry species. The molecular mechanisms underlying right ovary degeneration in embryonic chicken are currently not understood. Researchers are interested in the regulatory systems involved in the right ovary degeneration of the chick embryo. However, due to the scarcity of molecular data and computational resources, few investigations have been conducted. Studies have shown that ovary degeneration is a complex physiological process (Yu *et al.,* 2017). Transcriptome analysis can give gene expression data, which can be utilized to identify expression profiles linked with phenotypes of interest, such as degeneration. In this study, differential gene expression analysis during $E12^{th}$ day performed using RNA sequencing (RNA−Seq) transcriptome data analysis and validated using RT-PCR (Gibson *et al.,* 1996).

5.1 EMBRYO SEXING AND FEMALE GONADAL DEVELOPMENT

In chicken, the male is homogametic (ZZ) and the female is heterogametic (ZW). The left and right sides of the female chicken embryos at embryonic 4.5 days initially shared the same gonadal morphological appearance (Li *et al.,* 2022), indicating there is no asymmetric growth at embryonic 4.5 days. The asymmetric development is noticed on the 7^{th} embryonic day of development, the left gonads of the female chicken embryos were much larger than the right gonads. As a result, by the 10th day of embryonic development, the left gonad was over 2.5 times bigger than the right gonad (Li *et al.,* 2022). Over the course of their development, the maximum increase in the ovarian mass of a chicken embryo occurred in between days 12 and 13 of incubation (Khokhlov *et al.,* 2020). Most genes were regulated between 8 and 12 days of incubation (Dunislawska *et al.,* 2020).

The present study was performed to know differentially expressed genes between left and right ovaries during E12 stage where the largest increase in ovarian weight is noticed Morphologically male and female chick embryos cannot be distinguished at $E12^{th}$ day of development. Only molecular techniques can determine the gender of chicken embryos. In the present study, SWIM primers (He *et al.,* 2019)

sequences present on chromosome W were used to identify sex of embryos (Fig. 4.37). In conclusion, SWIM sexing is quick PCR sexing methodology for chickens.

High-throughput RNA-sequencing techniques were used and identified 539 and 1046 genes that are significantly differentially expressed between the embryonic left and right chicken ovaries on 6th day and 10th day respectively (Yu *et al.,* 2017). In the current study, 9152 genes were revealed to be significantly differentially expressed between the embryonic left and right chicken ovaries on the 12th day. This analysis clearly shows the highest transition in right ovary regression happens during the embryonic 12th day.

5.2 TRANSCRIPTOME ANALYSIS FOR IDENTIFICATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS

The genome is generally fixed for all cells throughout their lifetime, the transcriptome displays genes that are expressed at a specific time and can change based on several environmental circumstances. To understand the right ovary regression, the dynamics underlying the transcriptome of the left and right ovary need to be studied.

The datasets (SRR4029464, SRR4029463, SRR4029462 and SRR4029461) from NCBI were utilized in the current work to analyze the differentially expressed genes in the left and right ovary of chicken during its embryonic 12th day to find the genes responsible for the regression of the right ovary during the embryonic phase. The obtained raw data was found to have a total of 72.7 million reads. The FastQC report was found to have satisfactory quality parameters and does not have any adapter sequences. A total of 66.2 million processed reads were obtained after filtration for its quality (Q30) using Trimmomatic. When the processed reads were aligned to the *Gallus gallus* reference genome (GRCg7b), 53.06 million reads were uniquely mapped resulting in a mapping percentage of 80.22. The total number of significantly up-regulated genes by a fold change above +2 in the right ovary during E12 was 479 and the significantly down-regulated genes by fold change below -2 in the right ovary during E12 were 227.

Histological and stereological examination explored changes that occur in numerous components of the developing left and regressing right ovaries in chicken. It has been noticed that chicken left ovary contains a cortex and a medulla, whereas the

right ovary only has a medulla (González-Morán 2011). Similarly, histological examination of the right and left ovaries in ostrich embryos was examined and discovered that the right ovary had a thin layer of germinal epithelium (Nabipour *et al.,* 2015).

In the present study, transcriptome analysis revealed CYP19A1 (cytochrome P450 / Aromatase) and ESR1(estrogen receptor 1) have been downregulated by a fold change of -1.84 and -2.6 in the right ovary, which plays a key role in the right ovary cortex and medulla development. However, downregulation of CYP19A1 and ESR1 gene in the right ovary leads to lower estrogen production in the right ovary and higher levels of estrogen in the left ovary resulting in the development of a functional left ovary. During chicken embryonic development (E10–E18), increased expression of FSHR (follicle stimulating hormone receptor) and CYP19A1 was found higher in the left ovary as compared to the right ovary, favouring left ovarian development and functionality (Shaikat *et al.,* 2018). The current transcriptome data analysis between E12 day right and left ovary shows a decreased expression of FSHR by fold change -2.05, and CYP19A1 by a found change of -1.84 in the right ovary, reducing right ovarian development and functionality. Experimentally induced female-to-male reversal with the aromatase inhibitor, Ovarian-specific BMP7 expression was reduced in the right ovary and BMP3 has more expression in the left during early and late embryonic stages (Hoshino *et al.,* 2005). In the current transcriptome analysis, BMP7 is down-regulated in the right ovary by a fold change of -1.2 and no significant differential expression was noticed with BMP3 during E12 day.

Three genes, SOX9, IGF1, and GAL, were expressed more on the left in later ovarian stages than on the right. In contrast, CYP11A1, FST, and PAX2 were expressed more on the right (Carré *et al.,* 2011). Transcriptome analysis of differentially expressed genes between left and right ovary during embryonic 12^{th} day reveals down-regulation of SOX9, IGF1, and CYP11A1 by a fold change of -1.2, -1.6 and -1.84 respectively and up-regulation of GAL and PAX2 by a fold change of +1.15 and +2.5 respectively in the right ovary. There is no significant difference in FST gene expression.

5.2.1 Enrichment Analysis

In the present study, differentially expressed genes were mapped to Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for investigating biologically significant functional pathways underlying right ovary degeneration using gProfiler (Reimand *et al.*, 2007) and ShinyGO (Ge *et al.*, 2020).

Most up-regulated genes are involved in inhibitory actions or in negative regulation pathways (Fig 4.13) whereas down-regulated genes are involved in the reduction of the meiotic cycle Process (GO:1903046) and can be seen in Fig 4.16. Ancestor chart of the meiotic process for GO:1903046 shown in Fig 4.35, concluding TDRD12, SYCP3, SMC1B and DDX4 with a fold change of -3.7, -3.1, -3.6 and -3.5 are involved in retardation of meiotic cell cycle process in the right ovary.

Pathway Analysis of up-regulated, down-regulated and combined genes analysis was performed and visualised as the tree plot (Fig 4.15a, Fig.4.18a, Fig.4.24), pathway network (Fig 4.15b, Fig.4.18b, Fig.4.23) and bar-plot (Fig 4.16, Fig. 4.19, Fig.4.22). The up-regulated genes SERPINC1, FGG and APOH have a fold change of +4.8, +4.6 and +4.9 respectively favouring blood coagulation, and fibrin clot formation. Genes such as SERPINC1, TF, ITIH2, AHSG, FETUB, HRG, OVALY, SERPINA10, AMBP, SERPIND1 and KNG1 have fold changes +4.8, +4.7, +4.4, 4.8, +4.9, +4.7, -2.9, +4.2, +4.0, +3.9 and +3.6 were involved in Endopeptidase inhibitor activity, Peptidase inhibitor activity, Endopeptidase regulator activity and Peptidase regulator activity pathways. Down-regulated genes like ASZ1, TDRD12, TDRD9, DDX4, TDRD1 and MOV10L1 have fold changes of -2.93, -3.7, -3.9, -4.1, -3.3 and -3.5 are involved in the pathways like piRNA metabolic process, DNA alkylation, DNA methylation and DNA modification. DAZL, ASZ1, SYCP3, TDRD12, TEX14, TDRD9, SYCP2, DDX4, MOV10L1 are linked with retardation of pathways like Meiotic nuclear division, Meiotic cell cycle process, nuclear division, organelle fission, Cellular proc. involved in reproduction in a multicellular organism.

5.2.1.1 Blood Clotting Cascade Pathway

FGA, FGG and SERPINB2 with fold change +4.8, + 4.6 and -3.7 respectively are involved in the Blood Clotting Cascade Pathway (Fig. 4.25). FGA and FGG have upregulated genes in the right ovary which are involved in Fibrin Monomer formation, which further with help of clotting factor XIIIa produces Fibrin (Horan *et al.,* 2001). FGA, FGB, and FGG are the three polypeptide chains that makeup fibrinogen (Yu *et al.,* 2021). An increase in FGA, FGB and FGG indicates that the pathway is activated and has developed a hypercoagulable state (Yu *et al.,* 2021). SERPINB2 is down-regulated in the right ovary, which inhibits the conversion of fibrin into fibrin split products, ultimately increasing Fibrin which aid in blood clotting.

5.2.1.2 Peroxisome proliferator-activated receptors signalling pathway

FABP1, APOC3, and HMGCS2 have fold changes of +4.0, +4.0, and +3.9 are involved in the Peroxisome proliferator-activated receptors (PPAR) signalling pathway (Fig 4.26). PPARs are nuclear hormone receptors that are activated by fatty acids, which are involved in lipid oxidation and clearance of circulating or cellular lipids (Moraes *et al.,* 2006). The integrating pathway of the above-mentioned genes is highlighted in pink colour in Fig 4.26. FABP1 helps VLDL Chylomicrons to enter the cells and their transportation, HMGCS2 helps in ketogenesis and APOC3 helps in lipid transportation. All the above-mentioned 3 genes collectively take part in lipid metabolism. Adipocyte differentiation is encouraged by PPAR-gamma to improve blood glucose absorption. Cell growth and lipid oxidation are both regulated by PPAR-delta. Therefore, all PPAR subtypes are essential for controlling lipid metabolism throughout the body (Takahashi *et al.,* 2005).

5.2.1.3 Insulin signalling pathway

Enrichment analysis also revealed that SHC4 with fold change -2.9 is involved in the Insulin signalling pathway (KEGG:04910) (Fig 4.28). Insulin receptors (INSR) can upregulate SHC receptors (Pessin and Saltiel, 2000) which in turn help in cell proliferation, and differentiation and trigger protein synthesis. In contrast, the right ovary has down-regulated SHC4, indicating reduced insulin in the cells and reduced cell proliferation, differentiation and protein synthesis.

5.2.1.4 Ferroptosis pathway

TF (transferrin/ovotransferrin) up-regulated by a fold change of +4.7 in the right ovary is involved in the ferroptosis pathway (Fig 4.29). Ferroptosis is a nonapoptotic, peroxidation-driven form of controlled cell death that requires a lot of iron (Yang and Stockwell 2016). A controlled form of cell death known as ferroptosis is characterized by the formation of reactive oxygen species (ROS) from accumulated iron and lipid peroxidation. Up-regulation of TF increases the accumulation of iron (Fe^{3+}) in the endosome and enhances the formation of reactive oxygen species (ROS) which ultimately leads to cell death.

5.2.1.5 Cell apoptosis

Microtubules of the eukaryotic cytoskeleton are composed of a heterodimer of alpha and beta tubulin (Dutcher, 2001). TUBA3E (Tubulin Alpha-3E Chain) encodes alpha tubulin (Abuhatzira *et al.*, 2009) and is downregulated by a fold change of -3.3 in the right ovary, leading to disruption of microtube function and leading to cell Apoptosis (KEGG:04210) (Fig 4.30).

TERB1, ASZ1, TOM1L1, SMC1B, DAZL, and TDRD9 are down-regulated genes involved in nuclear division (GO:0000280) (Fig.4.13) causing reduced nuclear division. CPB2, APOH, and FGG are up-regulated genes involved in causing Fibrinolysis (GO:0042730) (Fig 4.32). FETUB, AHSG, SPIA4, SERPINC1, HRG, ITIH2, SERPINA10, AMBP, SERPIND1, KNG1, OVALY, SERPINB2 have upregulated genes involved in Endopeptidase inhibitory activity (GO:0004866) (Fig 4.33). CPB2, APOH, SERPINC1, FGA, FGG, SERPINA10, SERPIND1, PROC, and KNG1 are collectively causing Blood coagulation (GO:007596) (Fig 4.34). TERB1, ASZ1, DAZL, and TDRD9 collectively cause retardation in the Meiotic cell cycle process (GO:1903046) (Fig 4.35).

5.3 VALIDATION OF TRANSCRIPTOME DATA

To verify the accuracy of transcriptome sequencing data analysis, four candidate genes (two up-regulated and two downregulated) were selected for real-time quantitative polymerase chain reaction reverse transcription (RT-qPCR) validation. RT-qPCR SYBR green assay using specific primers was used to measure the gene expression of FETUB, LEAP2, DAZL and DDX4 genes in left and right ovaries. The qPCR conditions were standardized for target genes (FETUB, LEAP2, DAZL and

DDX4) and reference gene (β actin). The relative tissue expressions of FETUB, LEAP2, DAZL and DDX4 mRNA were normalized with the reference gene (β actin). The amplification specificity was confirmed by the derivative melt curve. The common comparative threshold method of analysis was used to study the relative gene expression of FETUB, LEAP2, DAZL and DDX4 genes in both left and right ovarian tissues.

On the 12th day of the chicken embryo, we examined the differential expression of the FETUB, LEAP2, DAZL, and DDX4 genes in the left and right ovaries of Kadaknath chicken embryos. Right ovaries showed the highest expression of FETUB and LEAP2, as well as the lowest expression of DAZL and DDX4, in accordance with the transcriptome data validating it. Relative Quantitation (RQ) using the Comparative C_T method(Livak and Schmittgen 2001) has been utilised for analysis. In comparison to the left ovaries (24.57±1.95), the right ovaries (22.61±0.73) showed higher levels of FETUB gene expression with a lower C_T. Similar to this, right ovaries showed higher LEAP2 gene expression levels with lower C_T (21.59±0.46) as compared to left ovaries (22.58±2.13). In comparison to the left ovaries (18.53±1.36), the right ovaries (22.67±1.13) showed lower levels of DAZL gene expression with greater C_T. Similar to the left ovaries (19.433±1.73), the right ovaries (22.63±0.68) showed lower levels of DDX4 gene expression with greater C_T. In the current RT-qPCR SYBR green assay gene expression was in accordance to the transcriptome data, the right ovaries had the highest expression of FETUB and LEAP2, and the lowest expression of DAZL and DDX4.

Thus, in the present study, 9152 genes were significantly differentially expressed between the embryonic left and right chicken ovaries on the $E12^{th}$ day. Enrichment analysis revealed differentially expressed genes are involved in blood clotting cascade, PARA signalling pathway, insulin signalling pathway, ferroptosis and apoptosis leading to right ovary degeneration. Validation of differentially expressed genes was performed using RT-PCR. Further more exploration is needed in deep understating of right ovary degeneration

SUMMARY

CHAPTER VI

SUMMARY

The chicken (*Gallus gallus domesticus*), which is a subspecies of the red jungle fowl, is a domesticated fowl, that originated in South East Asia. Chickens are one of the most prevalent and widely used domestic animals. India is the third-largest producer of eggs in the world, producing 7.22% of global egg production (20th Livestock Census). With an annual commercial egg production of over 95.17 billion eggs, the total number of eggs produced in our nation is approximately 114.38 billion. The principal site for egg production in chickens is the ovary, which has a long evolutionary history of variety and adaptability over time. The current study is to identify differentially expressed genes between functional left and regressing right ovary. The highest rise in the ovarian mass of a chicken embryo happened between days 12 and 13 of incubation, during the course of their growth (Khokhlov *et al.*, 2020). Most of the gene regulation occurred between days 8 and 12 of incubation (Dunislawska *et al.*, 2020). To understand the underlying transcripts involved in right ovary regression, transcriptome analyses have been performed in E12 embryos keeping left ovary transcripts as control and right ovary transcripts as treatment. The current study used the NCBI datasets (SRR4029464, SRR4029463, SRR4029462, and SRR4029461) to analyze the differentially expressed genes in the left and right ovary of chicken during its embryonic 12th day in order to identify the genes in charge of the right ovary's regression. It was discovered that there were 72.7 million readings in all of the collected raw data. The quality metrics in the FastQC report were deemed to be good, and there are no adapter sequences. After filtering for quality (Q30), a total of 66.2 million processed reads were obtained using Trimmomatic. 53.06 million of the processed reads were uniquely mapped to the *Gallus gallus* reference genome (GRCg7b), yielding a mapping percentage of 80.22. In the current study, 9152 genes were revealed to be significantly differentially expressed between the embryonic left and right chicken ovaries on the E12th day. Similarly, the total number of significantly up-regulated genes by a fold change above +2 in the right ovary during E12 was 479 and the significantly down-regulated genes by fold change below -2 in the right ovary during E12 were 227.

The genes CYP19A1 (cytochrome P450 / Aromatase) and ESR1(estrogen receptor 1) have been downregulated by a fold change of -1.84 and -2.6 in the right

ovary. Downregulation in the right ovary leads to lower estrogen production in the right ovary and higher levels of estrogen in the left ovary having a key role in the left ovary cortex and medulla development. The left ovary displayed higher levels of FSHR (follicle stimulating hormone receptor) and CYP19A1 expression during the development of the chicken embryo. This favoured the growth and functionality of the left ovary over the right ovary. In the right ovary, GAL and PAX2 are up-regulated by a fold change of 1.15 and +2.5, respectively, while SOX9, IGF1 and CYP11A1 are down-regulated by a fold change of -1.2, -1.6, and -1.84, respectively, during the embryonic 12th day. Up-regulated genes SERPINC1, FGG and APOH favouring blood coagulation and fibrin clot formation. Genes such as SERPIND1, TF, ITIH2, AHSG, FETUB, HRG, OVALY, SERPINA10, and AMBP are up-regulated genes involved in Endopeptidase inhibitor activity, Peptidase inhibitor activity, Endopeptidase regulator activity and Peptidase regulator activity pathways.

Down-regulated genes like ASZ1, TDRD12, TDRD9, DDX4, TDRD1 and MOV10L1 are involved in the pathways like piRNA metabolic process, DNA alkylation, DNA methylation and DNA modification. DAZL, ASZ1, SYCP3, TDRD12, TEX14, TDRD9, SYCP2, DDX4, MOV10L1 are linked with retardation of pathways like meiotic nuclear division, meiotic cell cycle process, nuclear division, Organelle fission.

In the right ovary, the upregulated genes FGA and FGG have a role in the synthesis of fibrin monomers, which are then converted into fibrin which aid in clotting, whereas SERPINB2 is down-regulated, which prevents fibrin from being converted into fibrin split products, ultimately increasing the amount of fibrin that promotes blood clotting.

The right ovary has down-regulated SHC4, indicating reduced insulin in the cells and reduced cell proliferation, differentiation and protein synthesis. Up-regulation of TF increases the accumulation of iron (Fe3+) in the endosome and enhances the formation of reactive oxygen species (ROS) which ultimately leads to cell death (ferroptosis). TUBA3E (Tubulin Alpha-3E Chain) is downregulated in the right ovary, which encodes alpha tubulin, leading to disruption of microtube function and leading to cell Apoptosis.

Reduced nuclear division is brought on by downregulated nuclear division-related genes such as TERB1, ASZ1, TOM1L1, SMC1B, DAZL, and TDRD9. Genes that are up-regulated and contribute to fibrinolysis include CPB2, APOH, and FGG. Blood coagulation is caused by CPB2, APOH, SERPINC1, FGA, FGG, SERPINA10, SERPIND1, PROC, and KNG1. Together, TERB1, ASZ1, DAZL, and TDRD9 slow down the meiotic cell cycle process.

On the 12^{th} day of the chicken embryo, we examined the differential expression of the FETUB, LEAP2, DAZL, and DDX4 genes in the left and right ovaries of Kadaknath chicken embryos. The relative tissue expressions of FETUB, LEAP2, DAZL and DDX4 mRNA were normalized with the reference gene (β actin). According to the transcriptome data, verifying it, the right ovaries displayed the highest expression of FETUB and LEAP2, as well as the lowest expression of DAZL and DDX4.

LITERATURE CITED

LITERATURE CITED

Abuhatzira L, Shemer R, Razin A. 2009. MeCP2 involvement in the regulation of neuronal α-tubulin production. *Human Molecular Genetics* **18**(8):1415-23.

Aduma N, Izumi H, Mizushima S and Kuroiwa A. 2019. Knockdown of DEAD-box helicase 4 (DDX4) decreases the number of germ cells in male and female chicken embryonic gonads. *Reproduction, Fertility and Development **31*** (5): 847-854.

Andrews JE, Smith CA, Sinclair AH. 1997 . Sites of Estrogen Receptor and Aromatase Expression in the Chicken Embryo. *General and Comparative Endocrinology*. **108**(2):182-90.

Anstaett O L, Brownlie J, Collins M E and Thomas C J. 2010. Validation of endogenous reference genes for RT-qPCR normalisation in bovine lymphoid cells (BL-3) infected with Bovine Viral Diarrhoea Virus (BVDV). *Veterinary immunology and immunopathology* **137** (3-4): 201-207.

Attia T H and Saeed M A. 2016. Next generation sequencing technologies: A short review. *Next Generat Sequenc and Applications* **1**(2).

BAHS 2020. Basic Animal Husbandry Statistics, Ministry of Agriculture, Government of India. Bagniski, E.S. 1973. Clinical Chemistry Acta.

Bansal G, Narta K and Teltumbade M R. 2018. Next-Generation sequencing: technology, advancements, and applications. In *Bioinformatics: Sequences, structures, phylogeny* (15-46). Springer, Singapore.

Bates D, Mächler M, Bolker B and Walker S. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv**:1406**:5823*.

Batut B, Beek M V D, Doyle M A and Soranzo N. 2021. RNA-Seq data analysis in galaxy. In *RNA Bioinformatics* (367-392). Humana, New York, NY.

Batut B, Beek MV, Doyle MA, Soranzo N. 2021. RNA-Seq data analysis in galaxy. *InRNA Bioinformatics* (pp. 367-392). Humana, New York, NY.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15):2114-20.

Bruggeman V, Van As P and Decuypere E. 2002. Developmental endocrinology of the reproductive axis in the chicken embryo. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **131** (4): 839-846.

Bullard JH, Purdom E, Hansen KD, Dudoit S. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC bioinformatics* **11**(1):1-3.

Carlon N and Stahl A. 1985. Origin of the somatic components in chick embryonic gonads. *Archives d'anatomie microscopique et de morphologie experimentale* **74** (1): 52-59.

Carré G A, Couty I, Hennequet-Antier C and Govoroun M S. 2011. Gene expression profiling reveals new potential players of gonad differentiation in the chicken embryo. *PLoS One* **6** (9): e23959.

Champion L R. 1955. A case of paired oviducts in the chicken. *Poultry Science* **34** (1): 184-186.

Chen X, Song P, Xia J, Guo J, Shi Y, Zhong Y and Li M. 2021. Evolutionarily conserved boule and dazl identify germ cells of Coilia nasus. *Aquaculture and Fisheries*.

Clinton M, Haines L, Belloir B and McBride D. 2001. Sexing chick embryos: a rapid and simple protocol. *British Poultry Science* **42** (1): 134-138.

Clinton M, Nandi S, Zhao D, Olson S, Peterson P, Burdon T and McBride D. 2016. Real-time sexing of chicken embryos and compatibility with in ovo protocols. *Sexual Development* **10** (4): 210-216.

Cock P J, Fields C J, Goto N, Heuer M L and Rice P M. 2010. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* **38 (**6): 1767-1771.

Dietzel E, Floehr J and Jahnen-Dechent W. 2016. The biological role of fetuin-B in female reproduction. *Annals of Reproductive Medicine and Treatment* **1** (1): 1003.

Ding H, Lin Y, Zhang T, Chen L, Zhang G, Wang J, ...and Dai G. 2021. Transcriptome Analysis of Differentially Expressed mRNA Related to Pigeon Muscle Development. *Animals* **11** (8): 2311.

Dunislawska A, Szczerba A, Siwek M and Bednarczyk M. 2020. Dynamics of the transcriptome during chicken embryo development based on primordial germ cells. *BMC Research Notes* **13** (1): 1-6.

Dutcher SK. 2001. The tubulin fraternity: alpha to eta. *Current Opinion in Cell Biology* **13**(1):49-54.

Fan X C, Liu T L, Wang Y, Wu X M, Wang Y X, Lai P, ... and Zhao G H. 2020. Genome-wide analysis of differentially expressed profiles of mRNAs, lncRNAs and circRNAs in chickens during Eimeria necatrix infection. *Parasites and Vectors* **13** (1): 1-18.

FAO (2007). The State of the World's Animal Genetic Resources for Food and Agriculture,. FAO Rome.

FAO. 2019. FAO statistic database (http://www.faostat.fao.org).

FAO. 2021. FAO statistic database (http://www.faostat.fao.org).

Fox E J, Reid-Bayliss K S, Emond M J and Loeb L A. 2014. Accuracy of next generation sequencing platforms. *Next generation, sequencing and applications* **1**.

Frota I M, Leitão C C, Costa J J, Brito I R, van den Hurk R and Silva J R. 2011. Stability of housekeeping genes and expression of locally produced growth factors and hormone receptors in goat preantral follicles. *Zygote* **19 (**1): 71-83.

García-García G, Baux D, Faugère V, Moclyn M, Koenig M, Claustres M and Roux A F. 2016. Assessment of the latest NGS enrichment capture methods in clinical context. *Scientific Reports* **6** (1): 1-8.

Ge SX, Jung D, Yao R. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**(8):2628-9.

Gibson U E, Heid C A and Williams P M. 1996. A novel method for real time quantitative RT-PCR. *Genome Research* **6** (10): 995-1001.

González-Morán M G, González-Arenas A, Germán-Castelán, L and Camacho-Arroyo I. 2013. Changes in the content of sex steroid hormone receptors in the growing and regressing ovaries of Gallus domesticus during development. *General and Comparative Endocrinology* **189**: 51-58.

González-Morán M G. 2011. Histological and stereological changes in growing and regressing chicken ovaries during development. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* **294** (5): 893-904.

Grzegorzewska A K, Sechman A, Paczoska-Eliasiewicz H E and Rząsa J. 2009. The expression of pituitary FSHβ and LHβ mRNA and gonadal FSH and LH receptor mRNA in the chicken embryo. *Reproductive Biology* **9** (3): 253-269.

Guioli S, Lovell-Badge R. 2007. PITX2 controls asymmetric gonadal development in both sexes of the chick and can rescue the degeneration of the right ovary. *Development* **134** (23): 4199–4208

Guioli S, Nandi S, Zhao D, Burgess-Shannon J, Lovell-Badge R and Clinton M. 2014 Gonadal Asymmetry and Sex Determination in Birds.. *Sexual Development*.

Gullapalli RR, Desai KV, Santana-Santos L, Kant JA, Becich MJ. 2012. Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics. *Journal of pathology informatics* **3**(1):40.

Hamburger V and Hamilton H L. 1992. A series of normal stages in the development of the chick embryo. *Developmental Dynamics* **195** (4): 231-272.

Harvey E and Holmes E C. 2022. Diversity and evolution of the animal virome. *Nature Reviews Microbiology* **20** (6): 321-334.

Hassanzadeh B, Nabipour A, Behnam Rassouli M and Dehghani H. 2014. Morphological development of testes in ostrich (Struthio camelus) embryo. *Anatomical Science International* **89** (3): 129-139.

He L, Martins P, Huguenin J, Van TN, Manso T, Galindo T, Gregoire F, Catherinot L, Molina F, Espeut J. 2019. Simple, sensitive and robust chicken specific sexing assays, compliant with large scale analysis. *PloS one* **14**(3):e0213033.

Heid C A, Stevens J, Livak K J and Williams P M. 1996. Real time quantitative PCR. *Genome Research* **6** (10): 986-994.

Horan JT, Francis CW. 2001. Fibrin degradation products, fibrin monomer and soluble fibrin in disseminated intravascular coagulation. *Seminars in Thrombosis and Hemostasis* **27** (06): 657-666

Hoshino A, Koide M, Ono T, Yasugi S. 2005 . Sex-specific and left-right asymmetric expression pattern of Bmp7 in the gonad of normal and sex-reversed chicken embryos. *Development, Growth & Differentiation* **47**(2):65-74.

Intarapat S and Stern C D. 2013. Sexually dimorphic and sex-independent left-right asymmetries in chicken embryonic gonads. *PLoS One* **8** (7): e69893.

Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Research* **48**(W1):W395-402.

Johnson A L. 2015. Reproduction in the female. In *Sturkie's avian physiology* (635-665). Academic Press.

Kanakachari M, Rahman H, Chatterjee R N and Bhattacharya T K. 2022. Signature of Indian native chicken breeds: a perspective. *World's Poultry Science Journal* **78** (2): 421-445.

Kang H J, Kawasawa Y I, Cheng F, Zhu Y, Xu X, Li M, ... and Šestan N. 2011. Spatio-temporal transcriptome of the human brain. *Nature* **478** (7370): 483-489.

Kheirabadi M, Nabipour A, Behnam Rassouli M, Dehghani H. 2014. Asymmetrically developed ovaries in embryo of ostrich (Struthio camelus). یازدهمین کنگره سراسری علوم تشریحی ایران.

Khokhlov RY, Kuznetsov SI. 2020. Growth of the chicken embryo and its reproductive organs. *Volga Region Farmland*. **3**(3):73-6.

Kim D and Langmead B. 2015. Salzberg SLHISAT. A fast spliced aligner with low memory requirements. *Nat Methods* **12** (4): 357-60.

Knief C. 2014. Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in plant science* **5**: 216.

Kubota K, Sato F, Aramaki S, Soh T, Yamauchi N and Hattori M A. 2007. Ubiquitous expression of myostatin in chicken embryonic tissues: its high expression in testis and ovary. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology* **148** (3): 550-555.

Lee P D, Sladek R, Greenwood C M and Hudson T J. 2002. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Research* **12** (2): 292-297.

Leinonen R, Sugawara H, Shumway M and International Nucleotide Sequence Database Collaboration. (2010). The sequence read archive. *Nucleic Acids research* **39**: D19-D21.

Li J, Sun C, Zheng J, Li J, Yi G and Yang N. 2022. Time-course transcriptional and chromatin accessibility profiling reveals genes associated with asymmetrical gonadal development in chicken embryos. *Frontiers in Cell and Developmental biology* 405.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**(7):923-30.

Liu X, Zhao J, Xue L, Zhao T, Ding W, Han Y and Ye H. 2022. A comparison of transcriptome analysis methods with reference genome. *BMC Genomics* **23** (1): 1-15.

Livak K J and Schmittgen T D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. M*ethods* **25** (4): 402-408.

Love M I, Anders S and Huber W. 2017. Analyzing RNA-seq data with DESeq2. *Bioconductor* **2**: 1-63.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**(12):1-21.

Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. 2017. Transcriptomics technologies. *PLoS Computational Biology*. **13**(5):e1005457.

Lu X, Huang L, Huang Z, Feng D, Clark R J and Chen C. 2021. LEAP-2: an emerging endogenous ghrelin receptor antagonist in the pathophysiology of obesity. *Frontiers in Endocrinology* **12**.

Malinen E, Kassinen A, Rinttila T and Palva A. 2003. Comparison of real-time PCR with SYBR Green I or 5′-nuclease assays and dot-blot hybridization with rDNA-targeted oligonucleotide probes in quantification of selected faecal bacteria. *Microbiology* **149** (1): 269-277.

Mardis E R. 2011. A decade's perspective on DNA sequencing technology. *Nature* **470** (7333): 198-203.

Mardis E R. 2013. Next-generation sequencing platforms. *Annual Review of Analytical Chemistry* ***6*** (1): 287-303.

Marguerat S and Bähler J. 2010. RNA-seq: from technology to biology. *Cellular and Molecular Life Sciences* **67** (4): 569-579.

Moore RJ, Doran TJ, Wise TG, Riddell S, Granger K, Crowley TM, Jenkins KA, Karpala AJ, Bean AG, Lowenthal JW. 2005. Chicken functional genomics: an overview. *Australian Journal of Experimental Agriculture* **45**(8):749-56.

Moraes L A, Piqueras L and Bishop-Bailey D. 2006. Peroxisome proliferator-activated receptors and inflammation. *Pharmacology and Therapeutics* **110** (3): 371-385.

Morrison TB, Weis JJ, Wittwer CT. 1998. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques* **24**(6):954-8.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**(7):621-8.

Musich R J. 2020. *A Recent (2020) Comparative Analysis of Genome Aligners Shows HISAT2 and BWA are Among the Best Tools*. Rochester Institute of Technology.

Mutz KO, Heilkenbrinker A, Lönne M, Walter JG, Stahl F. 2013. Transcriptome analysis using next-generation sequencing. *Current opinion in biotechnology*. **24**(1):22-30.

Nabipour A, Behnam Rassouli M, Dehghani H. 2015. Morphological development of ovaries in ostrich (Struthio camelus) embryo. *Comparative Clinical Pathology* **24**(5):1185-91.

Nie R, Zheng X, Zhang W, Zhang B, Ling Y, Zhang H, Wu C. 2022. Morphological Characteristics and Transcriptome Landscapes of Chicken Follicles during Selective Development. *Animals* **12**(6):713.

Palmer S, Wiegand A P, Maldarelli F, Bazmi H, Mican J M, Polis M and Coffin J M. 2003. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *Journal of Clinical Microbiology* **41** (10): 4531-4536.

Pereira R, Oliveira J and Sousa M. 2020. Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. *Journal of Clinical Medicine* **9** (1): 132.

Pessin JE, Saltiel AR. 2000. Signaling pathways in insulin action: molecular targets of insulin resistance. *The Journal of clinical investigation* **106**(2):165-9.

Pohjanvirta R, Niittynen M, Lindén J, Boutros P C, Moffat I D and Okey A B. 2006. Evaluation of various housekeeping genes for their applicability for normalization of mRNA expression in dioxin-treated rats. *Chemico-Biological Interactions* **160** (2): 134-149.

Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, Bell SM, Combaret V, Puisieux A, Mighell AJ, Robinson PA. 2003. Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC biotechnology* **3**(1):1-3.

Rao X, Huang X, Zhou Z and Lin X. 2013. An improvement of the 2ˆ (–delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostatistics, Bioinformatics and Biomathematics* **3** (3): 71.

Reddy P C, Sinha I, Kelkar A, Habib F, Pradhan S J, Sukumar R and Galande S. 2015. Comparative sequence analyses of genome and transcriptome reveal novel transcripts and variants in the Asian elephant Elephas maximus. *Journal of Biosciences* **40 (**5): 891-907.

Reimand J, Kull M, Peterson H, Hansen J and Vilo J. 2007. g: Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research* **35**:W193-W200.

Rengaraj D, Zheng Y H, Kang K S, Park K J, Lee B R, Lee S I, ... and Han J Y. 2010. Conserved expression pattern of chicken DAZL in primordial germ cells and germ-line cells. *Theriogenology* **74** (5): 765-776.

Rodler D, Stein K and Korbel R. 2015. Observations on the right ovary of birds of prey: a histological and immunohistochemical study. *Anatomia, Histologia, Embryologia* **44** (3): 168-177.

Romanowski T, Markiewicz A, Bednarz N and Bielawski K P. 2007. Housekeeping genes as a reference in quantitative real-time RT-PCR. *Postepy Higieny i Medycyny Doswiadczalnej (Online)* **61:** 500-510.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. *Cold spring harbor laboratory press*.

Scheider J, Afonso-Grunz F, Hoffmeier K, Horres R, Groher F, Rycak L, ... and Winter, P. (2014). Gene expression of chicken gonads is sex-and side-specific. *Sexual Development* **8** (4): 178-191.

Sell J. 1959. Incidence of persistent right oviducts in the chicken. *Poultry Science* **38**(1):33-5.

Shaikat A H, Namekawa S, Ahmadi S, Takeda M. and Ohkubo T. 2018. Gene expression profiling in embryonic chicken ovary during asymmetric development. *Animal Science Journal* **89** (4): 688-694.

Shen M, Wu P, Li T, Wu P, Chen F, Chen L, ... and Zhang G. 2020. Transcriptome analysis of circRNA and mRNA in Theca cells during follicular development in chickens. *Genes* **11** (5): 489.

Slatko B E, Gardner A F and Ausubel F M. 2018. Overview of next-generation sequencing technologies. *Current Protocols in Molecular Biology* **122** (1): e59.

Smith CA. Roeszler K N, Ohnesorg T, Cummins D M, Farlie P G, Doran T J and Sinclair A H. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* **461** (7261): 267-271.

Sturkie P D and Mueller W J. 1976. Reproduction in the female and egg production. In *Avian physiology* (302-330). Springer, Berlin, Heidelberg.

Su S, Dwyer D M, Miska K B, Fetterer R H, Jenkins M C ad Wong E A. 2017. Expression of host defense peptides in the intestine of Eimeria-challenged chickens. *Poultry Science* **96** (7): 2421-2427.

Sun T, Xiao C, Deng J, Yang Z, Zou L, Du W, ... and Yang X. 2021. Transcriptome analysis reveals key genes and pathways associated with egg production in Nandan-Yao domestic chicken. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **40**: 100889.

Sutherland D A T, Honaker C F and Siegel P B. 2017. Dual functioning ovaries and atresia in chickens. Is it a coincidence? *Poultry Science* **96** (10): 3763-3767.

Tajadini M, Panjehpour M, Javanmard SH. 2014. Comparison of SYBR Green and TaqMan methods in quantitative real-time polymerase chain reaction analysis of four adenosine receptor subtypes. *Advanced biomedical research* **3**.

Takahashi N, Goto T, Kusudo T, Moriyama T and Kawada T. 2005. The structures and functions of peroxisome proliferator-activated receptors (PPARs). *Nihon rinsho. Japanese Journal of Clinical Medicine* **63 (**4): 557-564.

Tan M P, Wong L L, Razali S A, Afiqah-Aleng N, Mohd Nor S A, Sung,Y Y, ... and Danish-Daniel M. 2019. Applications of next-generation sequencing technologies and computational tools in molecular evolution and aquatic animals conservation studies: A short review. *Evolutionary Bioinformatics* **15:** 1176934319892284.

Tanaka R, Izumi H and Kuroiwa A. 2017. Androgens and androgen receptor signaling contribute to ovarian development in the chicken embryo. *Molecular and Cellular Endocrinology* **443**: 114-120.

Tichopad A, Dilger M, Schwarz G and Pfaffl M W. 2003. Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Research* **31** (20): e122-e122.

Van Dijk E L, Auger H, Jaszczyszyn Y and Thermes C. 2014. Ten years of next-generation sequencing technology. *Trends in Genetics* **30** (9): 418-426.

Van Limborgh J, Van Faassen F. 1960. The asymmetry of the gonads in duck embryos experimentally turned on their right side. *Acta morphologica Neerlando-Scandinavica* **3**:79-91.

Villalpando I, Sanchez-Bringas G, Sanchez-Vargas I, Pedernera E and Villafan-Monroy H. 2000. The P450 aromatase (P450 arom) gene is asymmetrically expressed in a critical period for gonadal sexual differentiation in the chick. *General and Comparative Endocrinology* **117** (3) 325-334.

Von Ahsen N, Schutz E, Armstrong V W and Oellerich M. 1999. Rapid detection of prothrombotic mutations of prothrombin (G20210A), factor V (G1691A), and methylenetetrahydrofolate reductase (C677T) by real-time fluorescence PCR with the LightCycler. *Clinical Chemistry* **45** (5): 694-696.

Wan Z, Lu Y, Rui L, Yu X, Yang F, Tu C and Li Z. 2017. Gene expression profiling reveals potential players of left-right asymmetry in female chicken gonads. *International Journal of Molecular Sciences* **18** (6): 1299.

Wang T and Brown M J. 1999. mRNA quantification by real time TaqMan polymerase chain reaction: validation and comparison with RNase protection. *Analytical Biochemistry* **269** (1): 198-201.

Wen G. 2017. A simple process of RNA-sequence analyses by Hisat2, Htseq and DESeq2. In *Proceedings of the 2017 International Conference on Biomedical Engineering and Bioinformatics* (11-15).

Wolf J B. 2013. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Molecular Ecology Resources*: **13** (4): 559-572.

Wong M L and Medrano J F. 2005. Real-time PCR for mRNA quantitation. Biotechniques 39 (1): 75-85.

Yang W S and Stockwell B R. 2016. Ferroptosis: death by lipid peroxidation. *Trends in Cell Biology* **26** (3): 165-176.

Yousif A, Drou N, Rowe J, Khalfan M and Gunsalus K C. 2020. NASQAR: a web-based platform for high-throughput sequencing data analysis and visualization. *Bmc Bioinformatics* **21** (1): 1-14.

Yu C, Zhang T, Shi S, Wei T and Wang Q. 2021. Potential biomarkers: differentially expressed proteins of the extrinsic coagulation pathway in plasma samples from patients with depression. *Bioengineered* **12** (1): 6318-6331.

Yu J, Yan L, Chen Z, Li H, Ying S, Zhu H and Shi Z. 2017. Investigating right ovary degeneration in chick embryos by transcriptome sequencing. *Journal of Reproduction and Development*: 2016-134.

Zheng X, O'Connor J, Huchzermeyer F, Wang X, Wang Y, Wang M and Zhou Z. 2013. Preservation of ovarian follicles reveals early evolution of avian reproductive behaviour. *Nature*: **495** (7442): 507-511.

Zou X, Wang J, Qu H, Lv X H, Shu D M, Wang Y, ... and Liu D W. 2020. Comprehensive analysis of miRNAs, lncRNAs, and mRNAs reveals potential players of sexually dimorphic and left-right asymmetry in chicken gonad during gonadal differentiation. *Poultry Science* **99** (5): 2696-2707.

APPENDIX

APPENDIX I

COMPOSITION OF BUFFERS AND REAGENTS

1. Buffers and Reagents for the isolation of RNA

1.1 DEPC water

DEPC – 1mL

Double distilled water – up to 1000 mL

The ingredients were mixed overnight in conical flask using a magnetic stirrer, autoclaved at 121°C for 15 minutes and a pressure of 15 psi. The treated water was stored at room temperature until use.

2. REAGENTS FOR AGAROSE GEL ELECTROPHORESIS

2.1 10x TBE Electrophoresis buffer

Tris base - 108 g

Boric acid - 55 g

EDTA - 8.3 g

Distilled water up to 1000 mL

2.2 Agarose (1%)

Agarose- 1 g

TBE (1X) - 100 mL

2.3 Ethidium Bromide

Ethidium Bromide- 100 mg (10 mg/mL)

Distilled water 10 mL

2.4 6X Gel loading buffer

Bromophenol blue- 0.025 g

Xylene cyanol- 0.025g

Sucrose 4 g

Distilled water 10mL

3. PREPARATION AND CASTING OF 1% AGAROSE GEL

i. Weigh 1gm of agarose powder and mix it with 100 mL of 1X TBE buffer in a conical flask. Swirl to mix.

ii. Then heat the solution in the microwave. At 30 seconds intervals, remove the flask and swirl the contents to mix well. Repeat until the agarose has completely dissolved. Then remove the flask from the microwave and swirl gently.

iii. Add ethidium bromide to a concentration of 10 mg/mL to the melted agarose and mix well. Allow the agarose solution to cool to 55-60^{o}C.

iv. Place the gel tray into the casting apparatus. Tape the open edges of a gel tray to create a mold. Place an appropriate comb into the gel mold to create the wells.

v. Pour the molten agarose into the gel mold. Allow the agarose to cool at room temperature. Remove the comb and place the gel in the gel tank.

APPENDIX II

S No.	EQUIPMENTS	MANUFACTURERS
1	Bench top centrifuge	Sigma, USA
2	Electric balance	Sartorius, Germany
3	Gel document system	Bio-Rad, USA
4	Gel drier	Bio-Rad, USA
5	Laminar air flow	Sanyo, Japan
6	Microcentrifuge	Tarsons, India
7	Micropipettes	Eppendorf, Germany
8	Micro tips	Tarsons, India
9	Spectrophotometer	EVO 300, USA
10	Thermal cycler	Himedia, Germany
11	Vortex	Genei, Bangalore
12	Real time PCR	Himedia, Germany
13	PCR machine	Himedia, Germany