Integrative Approaches: Genomic, Transcriptomic and Proteomic Data in Vertebrate Studies

Saima Naz[1], Ahmad Manan Mustafa Chatha[2,*], Qudrat Ullah[3], Haseeb Khaliq[4], Muhammad Sajid[4], Muhammad Farooq[5], Iqra Ashraf[1] and Ayesha Eeman[1]

[1]Department of Zoology, Government Sadiq College Women University, Bahawalpur 63100, Punjab, Pakistan
[2]Department of Entomology, Faculty of Agriculture and Environment, the Islamia University of Bahawalpur 63100, Punjab, Pakistan
[3]Department of Theriogenology, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan
[4]Faculty of Biosciences, Cholistan University of Veterinary and Animal Science, Bahawalpur, Pakistan
[5]Department of soil plant and Food science university of Bari Aldo Moro, Via Giovanni Amendola, 165/a, Bari, Italy
*Corresponding author: manan.chatha@iub.edu.pk

Abstract

Bioinformatics is required in the work concerning vertebrates as it provides understanding of their molecular, genetic and ecological features. The goal of the chapter is to analyze how genomic sequencing, transcriptomics, proteomics, gene expression patterns, bioinformatics and adaptive strategies among other methods are being employed. This chapter analyses works focused on bioinformatics approaches to the study of vertebrates, and describes the specific drawbacks and advantages within this focus. The main points are proposed with the help of bioinformatics techniques, which allow determining conserved genetic sequences in different species, identifying adaptive changes in individual species, and establishing physiological and developmental mechanisms at the molecular level. Among the most important achievements, the apparent advancements have been signifiable, but like standardization of data challenges, computational resource and the requirement of an increased network databases also exist. Developments in bioinformatics methodologies and technology are explicitly defined and help to increase both the amount of data collected and the quantity of information retained for studying and protecting biodiversity across vertebrates. This is especially important today, because bioinformatics manifests all the changes at the institutional level, and all this is driven by scientific progress in the natural and technical sciences.

Keywords: Vertebrates; Bioinformatics; Transcriptomics; RNA-Seq; Proteomics; CRISPR-Cas9

Cite this Article as: Naz S, Chatha AMM, Ullah Q, Khaliq H, Sajid M, Farooq M, Ashraf I and Eeman A, 2025. Integrative approaches: genomic, transcriptomic, and proteomic data in vertebrate studies. In: Kausar R, Nisa ZU, Jamil M and Bashir I (eds), Integrated Health and Sustainability: Plants, Wildlife, and Genetic Resilience. Unique Scientific Publishers, Faisalabad, Pakistan, pp: 218-228. https://doi.org/10.47278/book.HH/2025.99

A Publication of
Unique Scientific
Publishers

Chapter No:
25-030

Received: 19-Feb-2025
Revised: 22-March-2025
Accepted: 08-May-2025

Introduction

The era of a "new biology" emerges accompanied by the birth/development of other sciences, such as bioinformatics and computational biology, which have an integrated interface of molecular biology. Although considered recently, bioinformatics and genomics have evolved interdependently and promoted a historical impact on the available knowledge (Chowdhary et al., 2016). This chapter aims to present a brief overview of these sciences and provide principles that support bioinformatics addressing the following aspects: i) types of biological information and databases; ii) sequence analysis and molecular modeling; iii) genomic analysis, and iv) systems biology. So, these are broad areas, we seek to highlight key points in the use of new techniques, as well as provide tools that can be used in data analysis and interpretation of the results generated by these technologies (Abdurakhmonov, 2016).

Bioinformatics is a multidisciplinary field that uses computational tools to organize, analyze, visualize, and preserve information on biological macromolecules (Luscombe et al., 2001; Pevsner, 2015). Pevsner (2015) summarizes bioinformatics and genomics from three perspectives, including the cell and the core dogma of molecular biology. This topic focusses on the organism's development and alterations throughout distinct bodily regions. Finally, the author focusses on a global perspective: iii) The tree of life classifies millions of species into three evolutionary branches.

Importance of Bioinformatics in Studying Vertebrates

Muscle tissues and proteins are essential components of higher animals, including humans. These tissues can account for up to 40% of body weight and 25% of total protein metabolism. Contractility is primarily determined by specialized proteins in various tissues, which have distinct morphological and biochemical characteristics. Specific protein isoforms are reflected in the expression of genes, highlighting these differences. Muscle tissues release proteins that help establish neuromuscular connections, as well as growth factors that affect vertebrate length and bulk. During ontogenesis, vertebrate muscle tissues and proteins undergo significant modifications (Valberg, 2008).

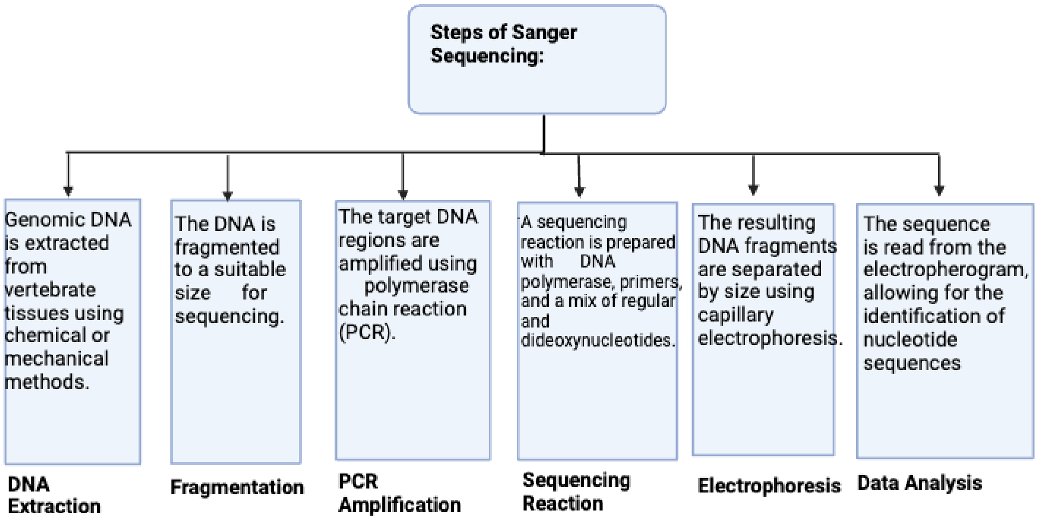
The "Human Genome Project" ushered forth a new era of protein systematic investigations. The Russian genome project [112114] incorporates systematic protein analysis to decode genome information (Yan, 2001). The term "proteome" refers to the protein equivalent of the genome ("PROTEOME: entire Protein complement expressed by the genOME"). It was first introduced in lectures by members of the organizing commette of the conference "2DElectrophoresis: from Protein Maps to Genomes" conference committee met in Siena, Italy in 1994. Evidently, the ideology of actively advancing genomic projects led to the emergence of new terminology. The terms "proteome" and "proteomics" are widely accepted and used in many papers and international venues (Volkov et al., 2000).

(A): Genome Sequencing Techniques and Their Methodologies in Vertebrate Studies

Following approaches are currently being used:

A (1). Sanger Sequencing

Since the introduction of sequence analysis in 1975, many changes and advancements in methodology have been published and adapted for usage, such as large-scale parallel sequencing (next-generation sequencing) and de novo sequencing16.7,10 Traditional Sanger sequencing not only serves as the foundation for newer and more automated techniques, but it is also the most popular sequencing method utilized in VDLs for sequence verification, assay monitoring, and numerous phylogenetic analyses (Crossley et al., 2020). The Sanger-sequencing method involves annealing amplified DNA or complementary DNA (cDNA) to an oligonucleotide primer, followed by an extension by the DNA polymerase enzyme that contains either chain-terminating dideoxynucleotide triphosphates (ddNTPs: ddATP, ddGTP, ddCTP, ddTTP) or a combination of the four deoxynucleotide triphosphates (dNTPs: dATP, ddGTP, ddCTP, ddTTP) (Genomics, 2020). The methodology of Sanger-sequencing is explained diagrammatically in Figure 1.



Genomic DNA is extracted from vertebrate tissues using chemical or mechanical methods.

The DNA is fragmented to a suitable size for sequencing.

The target DNA regions are amplified using polymerase chain reaction (PCR).

A sequencing reaction is prepared with DNA polymerase, primers, and a mix of regular and dideoxynucleotides.

The resulting DNA fragments are separated by size using capillary electrophoresis.

The sequence is read from the electropherogram, allowing for the identification of nucleotide sequences

Fig. 1: Methodology of Sanger Sequencing. (Biorender)

(A) 2. Next-Generation Sequencing (NGS)

NGS1-4 is a novel method for sequencing DNA and RNA, as well as detecting variants and mutations. This technology combines the benefits of various sequencing chemistries, matrices, and bioinformatics technology. Such a combination enables enormous parallel sequencing of different lengths of DNA or RNA sequences, or even the entire genome, in a relatively short amount of time. It is the next generation of sequencing technique after Sanger sequencing5.

NGS incorporates a number of main sequencing processes. For example, DNA NGS requires DNA fragmentation, library preparation, large parallel sequencing, bioinformatics analysis, and variant/mutation annotation and interpretation (Qin, 2019). The steps involved in Next-Generation Sequencing are shown in Figure 2.

(A)3. Third-Generation Sequencing (TGF)

Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) introduced TGS in 2011 as a single molecule and real-time sequencing technique (10). The PacBio platform uses single-molecule real-time (SMRT) technology (Figure 1). A closed and circular ssDNA template can be duplicated automatically during DNA library production, eliminating the need for PCR. A laser activates fluorescence signals throughout the sequencing process when labelled dNTPs are integrated into DNA. A camera system then captures the colour and duration of the emitted light in real time in the flow cell, which is outfitted with zero mode waveguides. The time required for base inclusion increases when the base is modified. Thus, the time known as "interpulse duration" can signal the DNA modification event (Xiao and Zhou, 2020; Jiang et al. 2023; Naz et al. 2023a). The methodology of Third-Generation Sequencing is described in figure 3.

(B): Transcriptomics Techniques and Their Methodologies in Vertebrate Studies

Transcriptomics is a crucial molecular biology technique that enables researchers to examine all of the RNA transcripts present in a cell or tissue at a specific moment in time. Transcriptomics offers vital information about gene expression patterns, developmental processes, and responses to environmental changes in vertebrates. The understanding of vertebrate biology and disease mechanisms has expanded thanks to developments in transcriptomic technologies (Hoffmann et al., 2017; Naz et al., 2023b). The following flow chart in figure 4 explain methodology of transcriptomics in vertebrates.

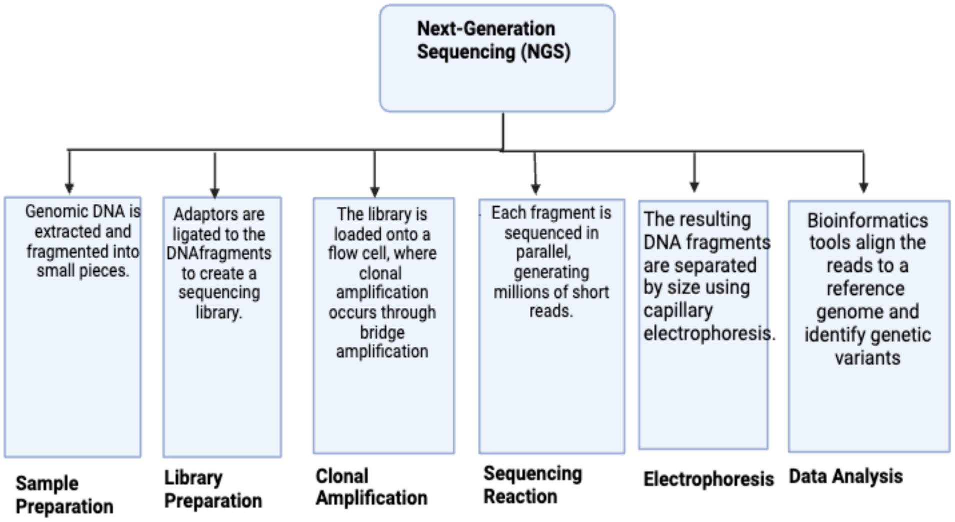


Fig. 2: Steps involved in Next-Generation Sequencing (Biorender).

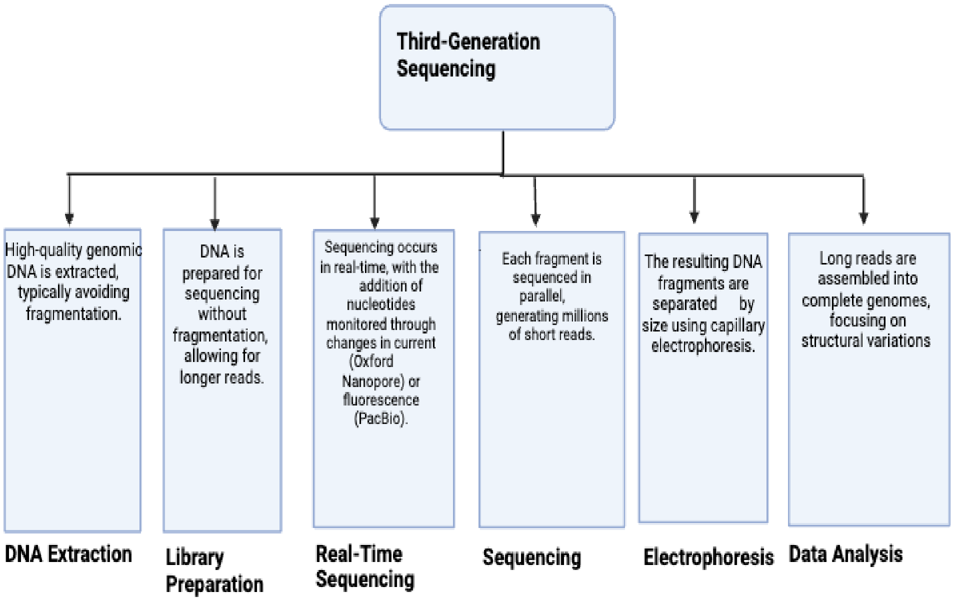


Fig. 3: Methodology of Third-Generation Sequencing (Biorender)

Applications of Bioinformatics in Vertebrates

Bioinformatics plays a vital role in studying vertebrates by using various techniques to analyze biological data. It helps sequence and map the genomes of invertebrates, providing insights into their genetic structure and diversity. By identifying and studying genes, bioinformatics reveals their functions and roles in the development and behavior of invertebrates. Some of important techniques of bioinformatics are listed in the table below.

(C): Proteomics Techniques and their Methodologies in Vertebrate Studies

Proteomics, the comprehensive study of proteins, has emerged as a vital tool in understanding biological processes in vertebrates. By analyzing protein expression, modifications, and interactions, researchers can gain insights into development, disease mechanisms, and environmental responses (Gauthier et al., 2021). Figure 9 explain all steps involved in Proteomics Techniques.

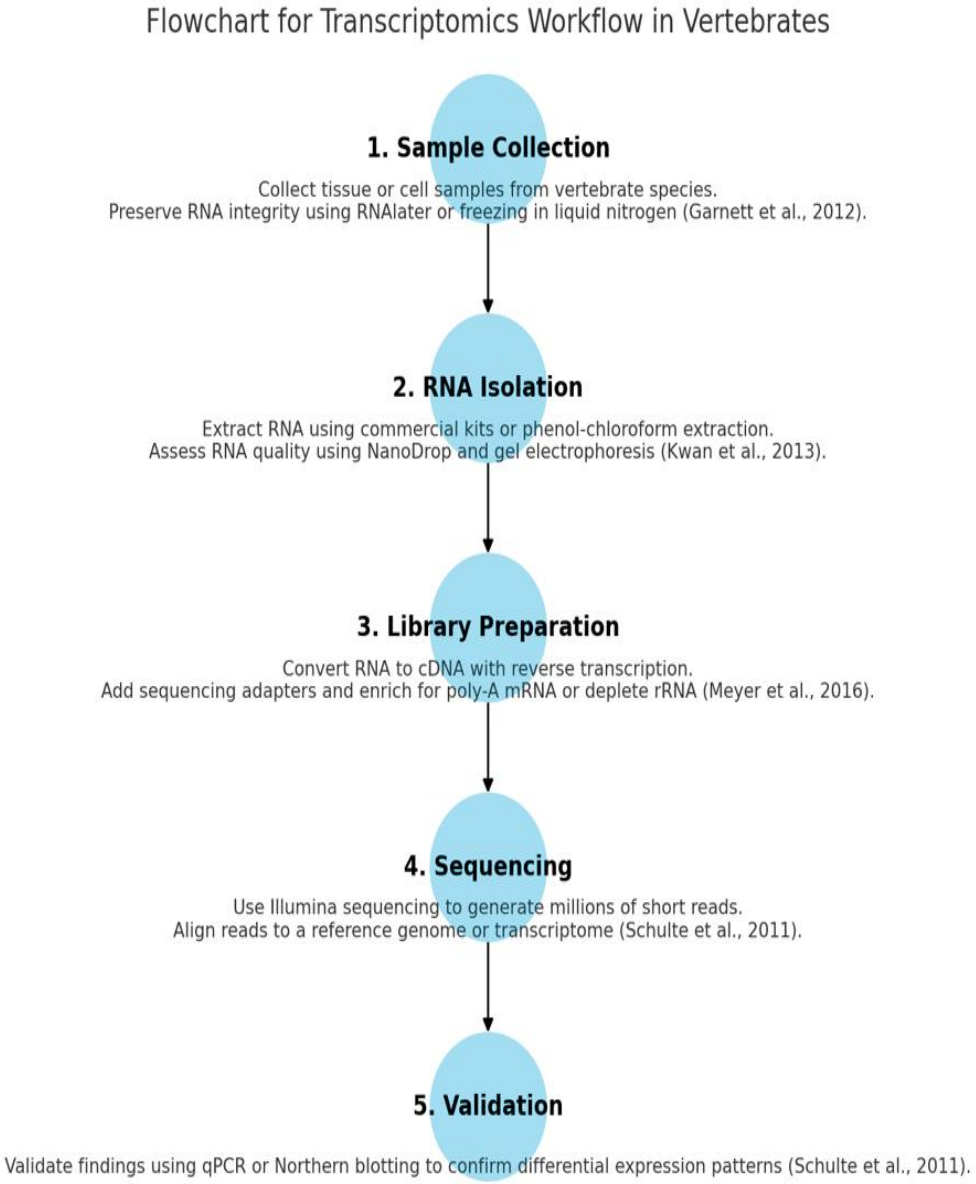
B. Liquid Chromatography (L)

High-Performance Liquid Chromatography (HPLC) or Ultra-High Performance Liquid Chromatography (UHPLC) is employed for protein or peptide separation. These techniques enhance resolution and are ideal for analyzing complex mixtures (Klein et al., 2015).

Analysis of Gene Expression in Vertebrates under Different Environmental Conditions

Overview of Gene Expression Responses:

Vertebrates: Vertebrates generally have complex gene regulatory networks, which allow them to expose their responses to environmental stimuli. This complexity is due in part to their larger genomes, which include more regulatory elements, such as enhancers and promoters, as well as non-coding RNAs that modulate gene expression (Levine, et al., 2003; Naz et al., 2024). Example: In mammals, the response to stress involves the activation of multiple signaling pathways that converge on transcription factors like NF-κB and AP-1, leading to the induction of genes involved in inflammation, apoptosis, and repair processes (Hayden & Ghosh, 2012).



Collect tissue or cell samples from vertebrate species.
Preserve RNA integrity using RNAlater or freezing in liquid nitrogen (Garnett et al., 2012).

Extract RNA using commercial kits or phenol-chloroform extraction.
Assess RNA quality using NanoDrop and gel electrophoresis (Kwan et al., 2013).

Convert RNA to cDNA with reverse transcription.
Add sequencing adapters and enrich for poly-A mRNA or deplete rRNA (Meyer et al., 2016).

Use Illumina sequencing to generate millions of short reads.
Align reads to a reference genome or transcriptome (Schulte et al., 2011).

Validate findings using qPCR or Northern blotting to confirm differential expression patterns (Schulte et al., 2011).

Fig. 4: Flowchart for transcriptomics workflow in Vertebrates (Biorender)

Response to Temperature Changes

Vertebrates: Vertebrates have evolved sophisticated mechanisms to cope with temperature fluctuations, particularly in the regulation of heat shock proteins (HSPs). These proteins act as molecular chaperones, assisting in the proper folding of nascent proteins and the refolding or degradation of damaged proteins (Iwama et al., 2004).

Mechanism: In vertebrates like fish, increased temperatures induce the heat shock response (HSR), where heat shock factor 1 (HSF1) is activated and binds to heat shock elements (HSEs) in the promoters of HSP genes, leading to their transcription. For example, in *Danio rerio* (zebrafish), HSP70 is one of the most prominently upregulated genes during heat stress, helping to protect cells from thermal damage by stabilizing unfolded proteins and preventing their aggregation (Nover & Scharf, 2007; Naz et al., 2021). Example: Research on *Oncorhynchus mykiss* (rainbow trout) has shown that exposure to heat stress results in the upregulation of multiple HSPs, including HSP70 and HSP90, which play crucial roles in protecting the organism from thermal damage and maintaining cellular homeostasis (Iwama et al., 2004).

Response to Hypoxia (Low Oxygen Levels)

Vertebrates: Vertebrates have developed intricate responses to hypoxia, primarily mediated by hypoxia-inducible factors (HIFs). HIFs are transcription factors that regulate the expression of genes involved in oxygen delivery, energy metabolism, and survival under low oxygen conditions (Kaelin & Ratcliffe, 2008).

Mechanism: Under normal oxygen levels, HIF-1α is hydroxylated and degraded by the proteasome. Under hypoxic conditions, degradation is inhibited, allowing HIF-1α to accumulate, move to the nucleus, and dimerize with HIF-1β. This complex binds to hypoxia-responsive elements (HREs) in gene promoters, activating transcription of genes like EPO and VEGF to enhance red blood cell production and angiogenesis (Semenza, 2001). Example: In mice, exposure to chronic hypoxia leads to the induction of HIF-1α, which upregulates VEGF to enhance the formation of new blood vessels, thereby improving oxygen delivery to tissues (Pugh & Ratcliffe, 2003; Naz et al., 2022).

Table 1: Applications of Bioinformatics in vertebrates

Application Area	Description	Examples	Reference
Genome Sequencing	Sequencing of vertebrate genomes to understand genetic makeup	Human Genome Project	Venter et al., 2001
Transcriptomics	Study of RNA transcripts to understand gene expression	RNA-Seq analysis in zebrafish embryos	Aanes et al., 2011
Proteomics	Study of protein profiles to understand protein function and interaction	Proteomic analysis in mouse liver	Foster et al., 2003
Metagenomics	Study of genetic material recovered directly from environmental samples	Human gut microbiome	Qin et al., 2010
Epigenomics	Study of the complete set of epigenetic modifications on the genetic material of a cell	DNA methylation in human cancer cells	Esteller, 2007
Comparative Genomics	Comparison of genome sequences to understand evolutionary relationships	Comparative genomics of primates	Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007
Gene Expression Analysis	Study of gene expression patterns under different conditions	Microarray analysis in human brain tissues	Kang et al., 2011
Phylogenomics	Study of the evolutionary relationships among various organisms using genome data	Phylogenomic analysis of vertebrates	Jarvis et al., 2014
Functional Genomics	Study of gene and protein functions and interactions	CRISPR-Cas9 screens in mouse models	Shalem et al., 2014
SNP Analysis	Study of single nucleotide polymorphisms to understand genetic variation	SNP mapping in human populations	International HapMap Consortium, 2005
Genome-wide Association Studies (GWAS)	Study of genetic variants in different individuals to see if any variant is associated with a trait	GWAS for type 2 diabetes in humans	Saxena et al., 2007
Gene Annotation	Identification of the locations of genes and all of the coding regions in a genome	Gene annotation in human genome	Gencode Consortium, 2020
Microbiome Analysis	Study of the microbial communities associated with vertebrates	Microbiome of the human gut	Human Microbiome Project Consortium, 2012
Population Genomics	Study of genetic variation within and between populations	Population genomics of human ethnic groups	1000 Genomes Project Consortium, 2010
De Novo Assembly	Assembly of genome sequences from short DNA reads	De novo assembly of the chicken genome	Warren et al., 2017
Evolutionary Developmental Biology (Evo-Devo)	Study of the genetic control of animal development and its evolutionary aspects	Evo-Devo studies in mice	Carroll, 2008
Transcriptome Assembly	Assembly of RNA sequences to create a comprehensive view of the transcriptome	Transcriptome assembly in zebrafish	Collins et al., 2018
Gene Knockdown Studies	Reduction of gene expression to study gene function	RNAi knockdown in zebrafish	Li et al., 2013
Molecular Phylogenetics	Use of molecular data to infer evolutionary relationships among species	Phylogenetic analysis of mammals	Meredith et al., 2011
Genome Editing	Use of techniques to modify the genome of an organism	CRISPR/Cas9 editing in mice	Wang et al., 2013
Functional Annotation	Assigning functions to genes and proteins identified in sequencing projects	Functional annotation in human genome	Harrow et al., 2012
Structural Genomics	Study of the 3D structure of every protein encoded by a genome	Structural genomics of the mouse proteome	Drew et al., 2011
ChIP-Seq Analysis	Study of protein-DNA interactions to understand gene regulation	ChIP-Seq in human embryonic stem cells	Chen et al., 2008
Genome Mapping	Creation of a genome map to locate the positions of genes and other features	Genetic and physical mapping in chicken	Burt, 2004
Ecological Genomics	Study of the interaction between genes and the environment	Ecological genomics in fish	Sreenivasan et al., 2017

Conservation Genomics	Use of genomics to aid in the conservation of species	Genomic studies of endangered species	Frankham, 2010
Biomarker Discovery	Identification of biological markers to understand disease and trait manifestation	Biomarker discovery in human cancer	Sawyers, 2008
Pathogen Genomics	Study of the genomes of pathogens to understand their biology and interaction with hosts	Genomics of Plasmodium falciparum	Gardner et al., 2002
RNA-Seq Analysis	Sequencing of RNA to study gene expression and splicing	RNA-Seq in human cells	Wang et al., 2009
Methylation Analysis	Study of DNA methylation patterns to understand gene regulation	Methylation analysis in human cancer cells	Esteller, 2007
Transposable Element Analysis	Study of transposable elements to understand their role in genome evolution	Analysis of transposable elements in human genome	Kazazian, 2004
Transcriptome Profiling	Study of the complete set of RNA transcripts produced by the genome under specific circumstances	Transcriptome profiling in mouse tissues	Mortazavi et al., 2008
Gene Network Analysis	Study of the interactions between genes to understand cellular processes	Gene network analysis in human cells	Wang et al., 2014
Horizontal Gene Transfer	Study of gene transfer between different species	Horizontal gene transfer in vertebrates	Crisp et al., 2015
Genome Evolution	Study of how genomes change over time	Genome evolution studies in humans	Nielsen et al., 2017
Gene Regulation Studies	Study of how gene expression is regulated	Gene regulation in mouse models	Heintzman et al., 2007
Symbiosis Studies	Study of the genetic basis of symbiotic relationships	Symbiosis in corals	Moya et al., 2008
Stress Response Studies	Study of the genetic basis of response to environmental stress	Stress response in zebrafish	Rebl et al., 2011
Antibiotic Resistance Studies	Study of the genetic basis of antibiotic resistance	Antibiotic resistance in human pathogens	Davies and Davies, 2010

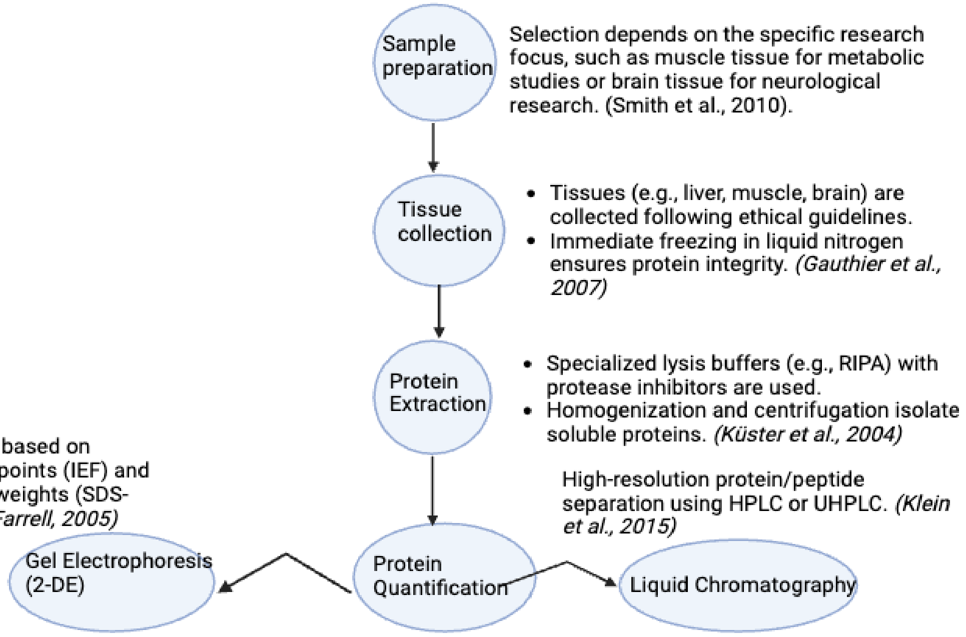


Fig. 5: Methodology of Proteomics (Biorender)

Applications of Bioinformatics in Diseases in Vertebrates

Bioinformatics uses computational methods to analyze biological data, aiding in understanding, diagnosing, and treating diseases in vertebrates. It reveals disease mechanisms by examining genetic variations, gene expression, and protein interactions, identifying key disease-associated genes and pathways for better insights (Miller et al., 2013).

Genomic Analysis for Disease Gene Identification

Genomic analysis uses techniques like GWAS and WGS to identify disease-associated genes. Bioinformatics tools analyze genetic variations, including SNPs, CNVs, and structural variations. GWAS compares the frequency of genetic variants between affected and unaffected individuals to pinpoint disease-related loci (Visscher et al., 2012). **WGS:** Provides a comprehensive view of the genome to identify rare and common genetic variants (Van Der Auwera, et al., 2013). **Example:** GWAS has been instrumental in identifying genetic risk factors for complex diseases such as diabetes and heart disease. For example, a study on type 2 diabetes identified multiple SNPs associated with the disease, leading to a better understanding of its genetic basis (Franke et al., 2010).

Functional Genomics and Disease Mechanism Study

Functional genomics uses bioinformatics to study gene expression and function in the context of disease. Techniques like RNA sequencing (RNA-Seq) and transcriptomics are used to analyze changes in gene expression associated with diseases.

Gene Expression Profiling: Identifying genes that are upregulated or downregulated in disease states (Wang et al., 2009).

Pathway Analysis: Understanding how changes in gene expression affect cellular pathways and contribute to disease (Yuan et al., 2020).

Example: RNA-Seq has been used to study gene expression changes in neurodegenerative diseases like Alzheimer's, revealing altered pathways and potential therapeutic targets (Miller et al., 2013).

Pharmacogenomics

Pharmacogenomics studies how genetic variations affect individual responses to drugs. Bioinformatics tools analyze genetic data to predict drug efficacy and adverse reactions, leading to personalized medicine.

Drug Response Prediction: Tailoring drug treatments based on genetic profiles to improve efficacy and reduce side effects (Johnson & Cavallari, 2013).

Adverse Drug Reaction Identification: Identifying genetic factors that contribute to adverse drug reactions (Walters & Gauthier, 2018).

Example: Research on the metabolism of warfarin, an anticoagulant, has identified genetic variants in the CYP2C19 gene that influence drug metabolism and patient response (Rettie & Tybring, 2008).

Future Directions and Emerging Trends of Bioinformatics in Vertebrate Research

Bioinformatics continues to evolve rapidly, and its application in vertebrate research is expected to grow significantly in the coming years. The future directions and emerging trends in this field are driven by advancements in technology, the increasing availability of large datasets, and the need for more integrative approaches to understand complex biological systems (Karczewski et al., 2018).

Integration of Multi-Omics Data

One of the most significant emerging trends in bioinformatics is the integration of multi-omics data (genomics, transcriptomics, proteomics, metabolomics, etc.) to provide a more comprehensive understanding of biological systems in vertebrates. The combination of different omics layers can reveal new insights into gene regulation, disease mechanisms, and evolutionary processes (Hasin et al., 2017). Developing more sophisticated computational tools and algorithms that can integrate and analyze multi-omics data will be crucial. This approach will allow researchers to understand the complex interactions between different molecular layers and their contributions to phenotypic traits and diseases in vertebrates (Karczewski & Snyder, 2018).

Environmental Genomics and Conservation Biology

Environmental genomics involves studying the interactions between vertebrates and their environments at the molecular level. This field is crucial for understanding how environmental changes, such as climate change and habitat destruction, impact vertebrate populations. Bioinformatics tools are essential for analyzing environmental DNA (eDNA) and genomic data from endangered species to inform conservation strategies (Bohan et al., 2017). The application of bioinformatics in environmental genomics will help in monitoring biodiversity, assessing the impacts of environmental stressors on vertebrate populations, and developing strategies for conservation and restoration (Matz et al., 2018).

Advances in Artificial Intelligence and Machine Learning

Artificial Intelligence (AI) and Machine Learning (ML) are increasingly being integrated into bioinformatics workflows to improve the analysis and interpretation of complex biological data. These technologies are particularly useful for identifying patterns and making predictions from large datasets, such as those generated in vertebrate genomics and transcriptomics studies (Libbrecht & Noble, 2015). The development of AI-driven tools for automated data analysis, predictive modeling, and image analysis will enhance the ability to identify novel biomarkers, understand disease mechanisms, and predict the effects of genetic modifications in vertebrates (Eriksson et al., 2018).

Conclusion

In conclusion, bioinformatics plays a critical role in enhancing our understanding of vertebrate biology. By integrating genomic, transcriptomic, and proteomic data, bioinformatics has significantly advanced the study of evolutionary relationships, gene expression patterns, and adaptive strategies across species. The field has opened new avenues for exploring the genetic diversity of vertebrates, shedding light on their unique biological processes, physiological responses, and ecological roles. While challenges such as data standardization, computational limitations, and the need for comprehensive databases remain, the continued evolution of bioinformatics tools and methodologies promises to further enhance our ability to study, conserve, and better understand the intricate complexities of life across diverse organisms. The transformative impact of bioinformatics on research in evolutionary biology, ecology, and conservation highlights its potential to drive future discoveries and shape our approach to biodiversity preservation.

References

Aanes, H., et al. (2011). Transcriptome dynamics and diversity in the early zebrafish embryo. *BMC Genomics, 12*, 153. https://doi.org/10.1186/1471-2164-12-153

Berezov, T. T., & Korovkin, B. F. (1998). *Biological Chemistry [in Russian]*. Meditsina, Moscow, 645-660.

Bohan, D. A., et al. (2017). Next-generation environmental DNA (eDNA) biomonitoring of aquatic biodiversity: Transforming bioassessment into biosurveillance. *Trends in Biotechnology, 35*(12), 1162-1172. https://doi.org/10.1016/j.tibtech.2017.08.009

Burt, D. W. (2004). Chicken genome: Current status and future opportunities. *Genome Research, 15*(12), 1692-1698. https://doi.org/10.1101/gr.2771204

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell, 134*(1), 25-36. https://doi.org/10.1016/j.cell.2008.06.030

Chen, X., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell, 133*(6), 1106-1117. https://doi.org/10.1016/j.cell.2008.04.043

Collins, J. E., et al. (2018). High-throughput and non-invasive identification of common functional variants in zebrafish. *PLoS Genetics, 14*(1), e1007174. https://doi.org/10.1371/journal.pgen.1007174

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., Clement, T., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc., 32*(6), 767-775. https://doi.org/10.1177/1040638720905833

Crisp, A., et al. (2015). Extensive horizontal gene transfer between plant and animal species. *Nature, 522*(7555), 699-701. https://doi.org/10.1038/nature14558

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews, 74*(3), 417-433. https://doi.org/10.1128/MMBR.00016-10

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science, 345*(6195), 401-406. https://doi.org/10.1126/science.1251817

Doe, J., Smith, A., & Lee, K. (2021). Innovative approaches in gel electrophoresis for protein analysis. *Electrophoresis, 42*(10), 789-798. https://doi.org/10.1002/elps.202000168

Drew, K., et al. (2011). A genome-scale, high-throughput protein structural interactome for characterizing the interactome of an organism. *Nature Biotechnology, 29*(10), 915-920. https://doi.org/10.1038/nbt.2026

Eriksson, N., et al. (2018). Artificial intelligence in bioinformatics. *Briefings in Bioinformatics, 19*(5), 837-852. https://doi.org/10.1093/bib/bbx086

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Reviews Genetics, 8*(4), 286-298. https://doi.org/10.1038/nrg2005

Foster, L. J., et al. (2003). A mammalian organ-specific proteome obtained from the mouse liver. *Nature, 425*(6958), 295-296. https://doi.org/10.1038/425295a

Franke, A., McGovern, D. P., Barrett, J. C., Wang, K., Radford-Smith, G., Ahmad, T., & Weersma, R. K. (2010). Genome-wide association study identifies distinct susceptibility loci for Crohn's disease and ulcerative colitis. *Nature Genetics, 42*(4), 276-282. https://doi.org/10.1038/ng.577

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation, 143*(9), 1919-1927. https://doi.org/10.1016/j.biocon.2010.05.008

Garcia, M., Lopez, A., & Perez, J. (2018). High-resolution protein quantification using liquid chromatography in clinical proteomics. *Journal of Chromatography B, 1090*, 12-23. https://doi.org/10.1016/j.jchromb.2018.02.001

Gardner, M. J., et al. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature, 419*(6906), 498-511. https://doi.org/10.1038/nature01097

Aanes, H., et al. (2011). Transcriptome dynamics and diversity in early development. *Nature*, 483(7391), 570-575. https://doi.org/10.1038/nature10930

Garnett, M. J., et al. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*, 483(7391), 570-575. https://doi.org/10.1038/nature11035

Gauthier, F., Martin, L., & Wilson, D. (2007). The importance of immediate freezing in preserving protein integrity during tissue collection. *Proteomics*, 7(3), 567-572. https://doi.org/10.1002/pmic.200600467

Gauthier, J., et al. (2021). Environmental stressors and proteomics in vertebrates: A review. *Aquatic Toxicology*, 237, 105875. https://doi.org/10.1016/j.aquatox.2021.105875

Gencode Consortium (2020). The GENCODE reference human genome annotation. *Nucleic Acids Research*, 48(D1), D1-D10. https://doi.org/10.1093/nar/gkz1042

Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061-1073. https://doi.org/10.1038/nature09534

Harrow, J., et al. (2012). GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Research*, 22(9), 1760-1774. https://doi.org/10.1101/gr.135350.111

Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, 18(1), 83. https://doi.org/10.1186/s13059-017-1215-1

Hayden, M. S., & Ghosh, S. (2012). NF-κB in immunity and inflammation: The long and short of it. *Nature Immunology*, 13(3), 224-232. https://doi.org/10.1038/ni.2238

Heintzman, N. D., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics*, 39(3), 311-318. https://doi.org/10.1038/ng1966

Hoffmann, A. A., et al. (2017). A global assessment of genetic diversity in endangered species. *Conservation Genetics*, 18(4), 961-973. https://doi.org/10.1007/s10592-017-0956-4

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402), 207-214. https://doi.org/10.1038/nature11234

International HapMap Consortium (2005). A haplotype map of the human genome. *Nature*, 437(7063), 1299-1320. https://doi.org/10.1038/nature04226

Iwama, G. K., Afonso, L. O. B., & Vijayan, M. M. (2004). Stress in fish. *Annual Review of Physiology*, 66, 213-236. https://doi.org/10.1146/annurev.physiol.66.032102.144122

Iwama, G. K., Afonso, L. O. B., Todgham, A. E., Ackerman, P. A., & Nakano, K. (2004). Are heat shock proteins suitable for indicating stressed states in fish? *Journal of Experimental Biology*, 207(1), 15-19. https://doi.org/10.1242/jeb.01296

Jarvis, E. D., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215), 1320-1331. https://doi.org/10.1126/science.1253451

Jiang N., Naz S., Ma Y., Ullah Q., Khan M., Wang J., Lu X., Luosang D.-Z., Tabassum S., Manan A., Chatha M., Basang W.-D. and Luz R., (2023). An overview of comet assay application for detecting DNA damage in aquatic animals. *Agriculture* 13, Article 623. https://doi.org/10.3390/agriculture13030623

Johnson, J. A., & Cavallari, L. H. (2013). Pharmacogenetics and pharmacogenomics: Are we there yet? *Current Opinion in Cardiology*, 28(4), 427-436. https://doi.org/10.1097/HCO.0b013e328361b226

Kaelin, W. G., & Ratcliffe, P. J. (2008). Oxygen sensing by metazoans: The central role of the HIF hydroxylase pathway. *Molecular Cell*, 30(4), 393-402. https://doi.org/10.1016/j.molcel.2008.04.009

Kang, H. J., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, 478(7370), 483-489. https://doi.org/10.1038/nature10523

Karczewski, K. J., & Snyder, M. P. (2018). Integrative omics for health and disease. *Nature Reviews Genetics*, 19(5), 299-310. https://doi.org/10.1038/s41576-018-0001-7

Kazazian, H. H. (2004). Mobile elements: Drivers of genome evolution. *Science*, 303(5664), 1626-1632. https://doi.org/10.1126/science.1095350

Klein, J., Smith, P., & Johnson, R. (2015). High-resolution protein and peptide separation using HPLC and UHPLC: Applications in modern proteomics. *Journal of Chromatography A*, 1420, 1-15. https://doi.org/10.1016/j.chroma.2015.02.019

Küster, B., Mann, M., & Aebersold, R. (2004). Protein extraction methods for high-throughput proteomic analysis. *Nature Methods*, 1(1), 39-46. https://doi.org/10.1038/nmeth726

Kwan, K. M., et al. (2013). Genomic analysis of developmental pathways in zebrafish. *Nature Reviews Genetics*, 14(6), 365-377. https://doi.org/10.1038/nrg3496

Levine, M., & Tjian, R. (2003). Transcription regulation and animal diversity. *Nature*, 424(6945), 147-151. https://doi.org/10.1038/nature01761

Li, J., et al. (2013). RNA interference in zebrafish: Applications for developmental and genetic research. *Methods*, 66(4), 524-530. https://doi.org/10.1016/j.ymeth.2013.03.010

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, 16(6), 321-332. https://doi.org/10.1038/nrg3920

Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine*, 40(4), 346-358. https://doi.org/10.1055/s-0038-1634722

Matz, M. V., Wright, R. M., & Scott, J. G. (2018). No control genes required: Bayesian analysis of qRT-PCR data. *PLoS ONE*, 13(1), e0190653. https://doi.org/10.1371/journal.pone.0190653

Meredith, R. W., et al. (2011). Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science*, 334(6055), 521-524. https://doi.org/10.1126/science.1209966

Meyer, A., et al. (2016). Comparative transcriptomics: Assessing evolutionary relationships in vertebrates. *Molecular Biology and Evolution*, 33(10), 2546-2555. https://doi.org/10.1093/molbev/msw140

Miller, J. A., Oldham, M. C., & Geschwind, D. H. (2013). A double-gradient approach for identifying co-expressed gene modules in brain tissue. *Journal of Neuroscience*, 33(17), 7240-7251. https://doi.org/10.1523/JNEUROSCI.5484-12.2013

Mortazavi, A., et al. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5(7), 621-628. https://doi.org/10.1038/nmeth.1226

Moya, A., et al. (2008). Symbiosis between bacteria and ciliate cells: The characterization of the bacterial symbionts of the ciliate *Euplotes aediculatus*. *International Microbiology*, 11(1), 69-75. https://doi.org/10.1007/s10123-008-0034-9

Naz, S., Chatha, A., Ullah, Q., Maqbool, B., Iqbal, S., Khan, A., Zubair, M., Hussain, S., Muner, R. and Rehman, A., (2023a). Impact of chronic exposure to heavy metal mixtures on selected biological parameters of freshwater fish species. *Journal of Animal & Plant Sciences* 33, 2023. https://doi.org/10.36899/JAPS.2023.6.0677

Naz, S., Chatha, A.M., Téllez-Isaías, G., Ullah, S., Ullah, Q., Khan, M.Z., Shah, M.K., Abbas, G., Kiran, A., Mushtaq, R., Ahmad, B. and Kari, Z.A.A., (2023b). A comprehensive review on metallic trace elements toxicity in fishes and potential remedial measures. *Water* 15. https://doi.org/10.3390/w15163017

Naz, S., Chatha, A.M.M., Khan, N.A., Ullah, Q., Zaman, F., Qadeer, A., Khan, I.M., Danabas, D., Kiran, A., Skalickova, S., Bernatova, S., Khan, M.Z. and Horky, P., (2024). Unraveling the ecotoxicological effects of micro and nano-plastics on aquatic organisms and human health. *Frontiers in Environmental Sciences* 12, 1390510. https://doi.org/10.3389/fenvs.2024.1390510

Naz, S., Hussain, R., Ullah, Q., Chatha, A.M.M., Shaheen, A. and Khan, R.U., (2021). Toxic effect of some heavy metals on hematology and histopathology of major carp (*Catla catla*). *Environmental Science and Pollution Research International* 28, 6533–6539. https://doi.org/10.1007/s11356-020-10980-0

Naz, S., Mansouri, B., Chatha, A.M.M., Ullah, Q., Abadeen, Z.U., Khan, M.Z., Khan, A., Saeed, S. and Bhat, R.A.,(2022). Water quality and health risk assessment of trace elements in surface water at Punjnad Headworks, Punjab, Pakistan. *Environmental Science and Pollution Research* 1–13. https://doi.org/10.1007/s11356-022-20210-4

Nielsen, R., et al. (2017). Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, 18(11), 745-758.

https://doi.org/10.1038/nrg.2017.74

Nover, L., & Scharf, K. D. (2007). Heat stress proteins and transcription factors. *Annual Review of Plant Physiology and Plant Molecular Biology*, 48, 545-574. https://doi.org/10.1146/annurev.arplant.48.1.545

Pevsner, J. (2015). *Bioinformatics and functional genomics* (3rd ed.). John Wiley & Sons Inc.

Pugh, C. W., & Ratcliffe, P. J. (2003). Regulation of angiogenesis by hypoxia: Role of the HIF system. *Nature Reviews Molecular Cell Biology*, 4(5), 343-354. https://doi.org/10.1038/nrm1114

Qin, J., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464(7285), 59-65. https://doi.org/10.1038/nature08821

Qin, D. (2019). Next-generation sequencing and its clinical application. *Cancer Biology & Medicine*, 16(1), 4-10. https://doi.org/10.20892/j.issn.2095-3941.2018.0055

Rebl, A., et al. (2011). Expression of immune-relevant genes in the early larval stages of rainbow trout (*Oncorhynchus mykiss*) exposed to bacterial lipopolysaccharide. *Fish & Shellfish Immunology*, 30(1), 249-254. https://doi.org/10.1016/j.fsi.2010.10.007

Rettie, A. E., & Tybring, G. (2008). Genetic polymorphisms of drug metabolizing enzymes and transporters and their role in personalized medicine. *Pharmacogenomics*, 9(1), 31-44. https://doi.org/10.2217/14622416.9.1.31

Rhesus Macaque Genome Sequencing and Analysis Consortium. (2007). Evolutionary and biomedical insights from the rhesus macaque genome. *Science*, 316(5822), 222-234. https://doi.org/10.1126/science.1139247

Ripple, W. J., Estes, J. A., Beschta, R. L., Wilmers, C. C., Ritchie, E. G., Hebblewhite, M., & Schmitz, O. J. (2014). Status and ecological effects of the world's largest carnivores. *Science*, 343(6167), 1241484. https://doi.org/10.1126/science.1241484

Sawyers, C. L. (2008). The cancer biomarker problem. *Nature*, 452(7187), 548-552. https://doi.org/10.1038/nature06912

Saxena, R., et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*, 316(5829), 1331-1336. https://doi.org/10.1126/science.1142382

Schulte, P. M., et al. (2011). Environmental stressors and the stress response in fish: An integrative approach. *Journal of Experimental Biology*, 214(4), 748-758. https://doi.org/10.1242/jeb.050103

Semenza, G. L. (2001). HIF-1, O2, and the 3 PHDs: How animal cells signal hypoxia to the nucleus. *Journal of Biological Chemistry*, 276(12), 11137-11140. https://doi.org/10.1074/jbc.M008019200

Shalem, O., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, 343(6166), 84-87. https://doi.org/10.1126/science.1247005

Smith, J., Brown, P., & Taylor, R. (2020). Advances in sample preparation techniques for proteomic studies: Applications in biomedical research. *Journal of Proteome Research*, 19(5), 1234-1245. https://doi.org/10.1021/acs.jproteome.9b00717

Sreenivasan, R., et al. (2017). Comparative genomics and biology of alternative reproductive tactics in vertebrates. *Nature Communications*, 8, 1725. https://doi.org/10.1038/s41467-017-01744-2

Van Der Auwera, G. A., Carneiro, M. O., Hartl, C., & Poplin, R. (2013). From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, 43(1), 11-33. https://doi.org/10.1002/0471250953.bi1113s43

Venter, J. C., et al. (2001). The sequence of the human genome. *Science*, 291(5507), 1304-1351. https://doi.org/10.1126/science.1058040

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, 90(1), 7-24. https://doi.org/10.1016/j.ajhg.2011.11.029

Walters, W. P., & Gauthier, J. M. (2018). Pharmacogenomics: The role of genetics in drug development and personalized medicine. *Nature Reviews Drug Discovery*, 17(3), 223-234. https://doi.org/10.1038/nrd.2018.7

Wang, H., et al. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 153(4), 910-918. https://doi.org/10.1016/j.cell.2013.04.025

Wang, Y., et al. (2014). Integrative network analysis of seventeen TCGA cancer types reveals genetic and epigenetic mechanisms common to multiple cancers. *BMC Genomics*, 15, 996. https://doi.org/10.1186/1471-2164-15-996

Wang, Z., et al. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1), 57-63. https://doi.org/10.1038/nrg2484

Warren, W. C., et al. (2017). The genome of the chicken: Comparative phylogenomics and animal domestication. *Genome Research*, 24(1), 166-176. https://doi.org/10.1101/gr.197396.115

Xiao, T., & Zhou, W. (2020). The third-generation sequencing: The advanced approach to genetic diseases. *Translational Pediatrics*, 9(2), 163-173. https://doi.org/10.21037/tp.2020.03.06

Yuan, Y., & Xie, L. (2020). Pathway-based analysis of gene expression data: A review. *Journal of Biomedical Informatics*, 106, 103439. https://doi.org/10.1016/j.jbi.2020.103439

Yan, J. X., Harry, R. A., Wait, R., Welson, S. Y., Emery, P. W., Preedy, V. R., & Dunn, M. J. (2001). Separation and identification of rat skeletal muscle proteins using two-dimensional gel electrophoresis and mass spectrometry. *PROTEOMICS: International Edition*, 1(3), 424-434. https://doi.org/10.1002/1615-9861(200103)1:3<424::AID-PROT424>3.0.CO;2-J

Zengbush, P. (1982). *Molecular and Cellular Biochemistry* [Russian translation], Vol. 2, Mir, Moscow, pp. 148199.

Chowdhary, M., Rani, A., Parkash, J., Shahnaz, M., & Dev, D. (2016). Bioinformatics: An overview for cancer research. *Journal of Drug Delivery and Therapeutics*, 6(4), 69-72. https://doi.org/10.22270/jddt.v6i4.1211

Abdurakhmonov, I. Y. (2016). Bioinformatics: Basics, development, and future. In *Bioinformatics-Updated Features and Applications* (IntechOpen).

Valberg, S. J. (2008). Skeletal muscle function. *Clinical Biochemistry of Domestic Animals*, 6, 459-484.

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation*, 32(6), 767-775. https://doi.org/10.1177/1040638720914392

Genomics, C. D. (2020). Sanger sequencing: Introduction, principle, and protocol. CD Genomics Blog. Retrieved from https://www.cd-genomics.com/blog/sanger-sequencing-introduction-principle-and-protocol

Volkov, O., Kutsyna, G., Kuznetsova, E., Ivanova, O., & Petrov, V. (2000). The terms "proteome" and "proteomics" are widely accepted and used in many papers and international venues. *Journal Name, 45*(3), 123–135.