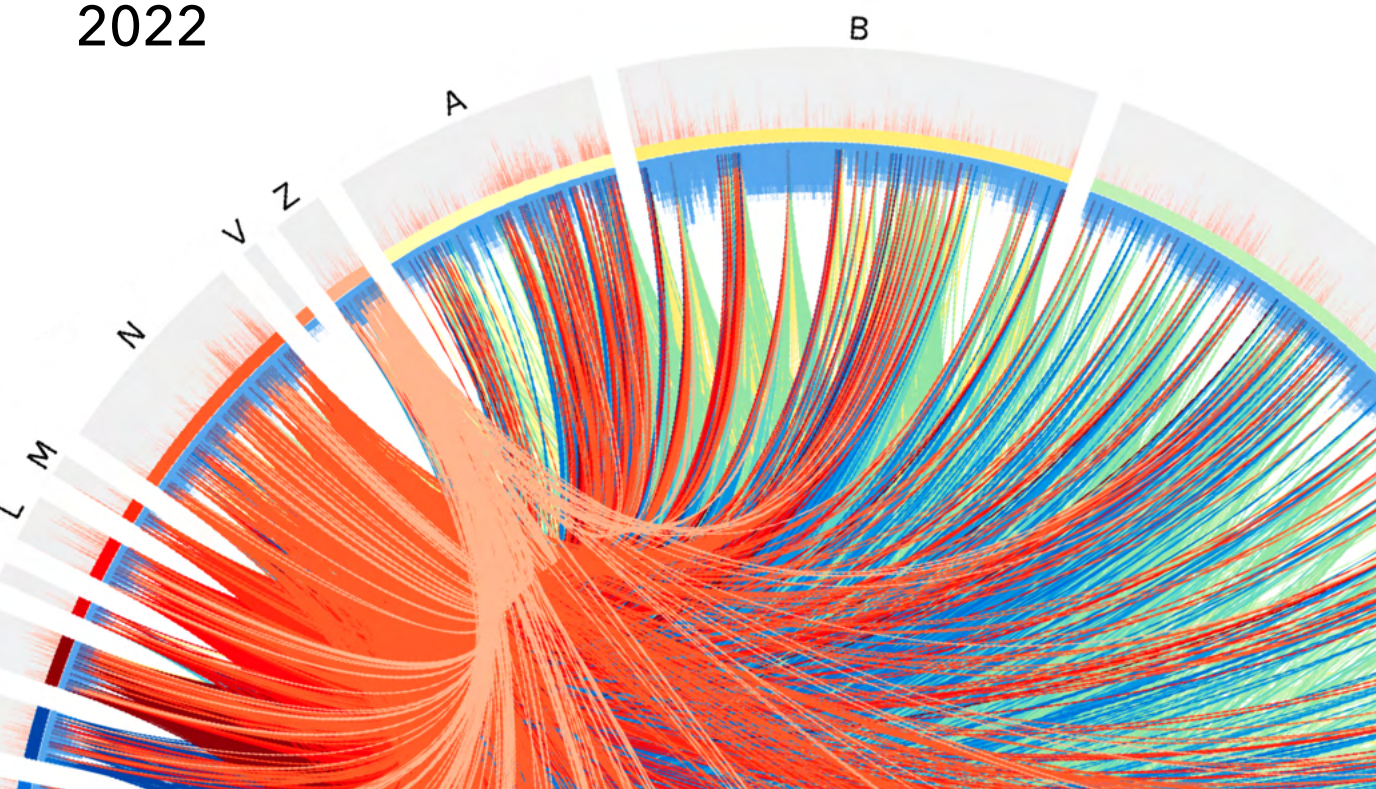
The Institute of Biology
Sri Lanka

# TOWARDS A PARADIGM SHIFT IN BIOLOGY

Thematic
Publication
2022



# Towards a Paradigm Shift in Biology

Published by

**The Institute of Biology Sri Lanka**

30th September 2022

**©The Institute of Biology Sri Lanka**
**September, 2022**

*All rights reserved. This book or any portion thereof may not be reproduced or used in any manner whatsoever without the written permission of the publisher except for the use of brief quotations in a book review.*

**Disclaimer:** The materials in this book have been provided by the authors and compiled by the Institute of Biology Sri Lanka (IOBSL). The contents and views expressed therein remain the responsibility of the named authors and do not necessarily reflect those of the copyright holder (IOBSL) and its members.

**Citation:** Name/s of chapter author/s (2022). 'Title of chapter', In: *Toward a Paradigm Shift in Biology*. Colombo; Institute of Biology Sri Lanka, pp. x-y

**Text by:** C.D. Dangalle, L. Abeywardhana, Y.W. Mallawarachchi, K.L.N. Perera, G. Galhena, W.R.P. Wijesinghe, D.V.S. Kaluthanthri, P.N. Dasanayaka, S.A.C.N. Perera, A. Bellanthudawa, M.D.K.L Gunathilaka, D. Halwatura

**Technical Editing:** R. Wimalasekera, G. Galhena

**Cover Page:** A portion of a Circos plot displaying the interconnectivity among different headings belonging to genomics-related PubMed-indexed articles reflecting a multitude of different areas of knowledge converging to conform a highly multidisciplinary corpus.
Courtesy of *M. Siqueiros-García, Jesús; Hernández-Lemus, Enrique; García-Herrera, Rodrigo; Robina-Galatas, Andrea (2015): Circos plot of MeSH-tree headings. PLOS ONE. Figure.*
*https://doi.org/10.1371/journal.pone.0092639.g004*

Designed by: Rovindu Kudaliyanage

**Printing and Binding:** The Press of the Open University of Sri Lanka

**ISBN:** 978-624-5998-02-9

Institute of Biology Sri Lanka (IOBSL)
'Vidya Mandiraya',
120/10, Wijerama Road, Colombo 07.
Email: iobslcouncil@gmail.com
Web: http://www.iobsl.org/

# Towards a Paradigm Shift in Biology

## Authors

*Chapter 1*
A new paradigm for insect classification: Using machine learning approaches
**C. D. Dangalle, L. Abeywardhana, Y. W. Mallawarachchi**

*Chapter 2*
Changing paradigms in forensic genetics
**K. L. N. Perera, G. Galhena**

*Chapter 3*
A journey into a microbial cell and beyond: The power of the lens in shifting microbiological paradigms
**W. R. P. Wijesinghe**

*Chapter 4*
Quantum Biology: A novel direction in biological sciences
**D. V. S. Kaluthanthri, P. N. Dasanayaka, S. A. C. N. Perera**

*Chapter 5*
Applications of Geographical Information Science (GIScience) in ecological and environmental research: Sri Lankan perspectives
**A. Bellanthudawa, M. D. K. L. Gunathilaka, D. Halwatura**

Published by

**The Institute of Biology Sri Lanka**

30th September 2022

# Preface

The shift from Newtonian physics to the theory of relativity, the introduction of Darwin's theory of evolution and Gregor Mendel's discovery on inheritance are some well-known paradigm shifts that have revolutionized scientific thinking. Together with more recent advancements made in understanding the complexity of organisms at both molecular and nano-levels, this wealth of knowledge has immensely contributed to the rapidly shifting landscapes observed in the field of biology.

The Institute of Biology, Sri Lanka (IOBSL), as the leading professional body for biologists in Sri Lanka recognizes the need to acknowledge these advances, to nurture the scientific community. Our thematic publication for the year 2022 titled "Towards a paradigm shift in Biology" is a tribute to these promising developments and compiles some of the highlights, cutting-edge tools, and novel concepts in biology, with the aim of inspiring scientists to be a part of this revolutionary phase.

This 'shift' is broad and encompasses an array of recent developments in genetics and genomics, synthetic and system biology, drug design and discovery, bionic designs, and machine learning to name a few. This book is but a blade of grass in this vast field of new innovations. Thus, we invite you to explore more to understand and make use of its enormous potential.

Nonetheless, we believe that the information provided in this book will be valuable to academics, researchers, teachers, and students in the field of biology and allied disciplines.

Our thanks go to the authors of the chapters, the reviewers, and the members of the council for their great support in producing this book.

Dr. R. Wimalasekera
Dr. G. Galhena
[Technical Editors]

30th September, 2022

# Reviewers

**Professor H. M. S. P. Madawala**
Department of Botany, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

**Professor G. R. Ranawaka**
Department of Zoology, Faculty of Natural Sciences, The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka

**Dr. P. C. Fernando**
Department of Plant Sciences, Faculty of Science, University of Colombo, Colombo, Sri Lanka

**Dr. R. J. K. U. Ranatunga**
Department of Chemistry, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

**Dr. D. Munasinghe**
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

# Acknowledgements

The Institute of Biology Sri Lanka would like to acknowledge the contribution of all those who have been involved in publishing this book. This book would not have become a reality without their generous support.

- We sincerely appreciate the contribution from chapter authors for their willingness and commitment to share their expertise by composing a chapter in this book.

- We gratefully acknowledge the efforts of chapter reviewers for their valuable contribution in improving quality, coherence, and content of the chapters.

- We are thankful to Mr. Rovindu Kudaliyanage for designing the cover page and title pages of the chapters.

# Contents

# Chapter 1

# A new paradigm for insect classification: Using machine learning approaches

C. D. Dangalle[1]*, L. Abeywardhana[2], Y. W. Mallawarachchi[2]

[1]Department of Zoology and Environment Sciences, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka.
[2]Department of Information Technology, Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka.
* Corresponding author: cddangalle@zoology.cmb.ac.lk

## Abstract

Insects are a highly diverse group of taxa inhabiting the Earth. Due to this high diversity and ecological and economic importance it is essential that they be accurately identified. Hitherto, morphological and molecular methods have been used with success to recognize the boundless variety of insect families, genera and species. However, certain drawbacks of these methods have noticeably affected the reliability of insect identification and the rapid accomplishment of the task. Therefore, many taxonomists and scientists are considering machine learning methods as a convenient option. In machine learning methods for automated insect identification, taxa are recognized using a computer-aided system that functions using insect data/ signals we introduce. These signals are preprocessed, subjected to feature extraction and classified into hierarchical levels for insect recognition. According to the insect being identified, various data such as acoustic signals, behavioural signals, and visual signals in the form of digital images can be used as input data, which in preprocessing are subsequently formatted to increase their quality and clarity. The most informative set of characters that define the insect are selected in feature extraction, and the taxon is assigned to a pre-specified taxonomic level based on the extracted features. At present, many countries have used machine learning techniques to identify insect species with great success. In Sri Lanka, studies using this approach has been conducted to identify insect pests of economic importance and biodiversity value.

## Keywords

*Artificial neural networks, Classification, Insects, Machine learning approach*

Insects form a large portion of the biological diversity of our planet and their rapid and reliable identification is important in many contexts. Accurate identification of insects is necessary for humans in the detection of pests of agricultural crops, pests of silviculture, recognition of disease vectors and the disclosure of invasive species (Eidt 1995; Valan et al., 2019). Further, reliable identification is necessary for the insects as well, in determining their diversity, phylogenetic patterns and evolutionary relationships which will be important in their conservation (Tahir et al., 2018). Hitherto, insect identification has been based upon morphological diagnosis and molecular analysis, which has been effective in providing us with the much needed facts and knowledge in insect classification. However, both approaches have experienced many difficulties and drawbacks that have ensued in challenges in insect classification.

## Morphological identification of insects and its drawbacks

Different insect species exhibit distinguishable morphological features which are utilized in taxonomic keys to identify them. However, morphological identification requires experienced taxonomists who are often in short demand, especially for groups that are not showy and attractive (Chan et al., 2014; Valan et al., 2019). Thus, there is a clear bias of focus on particular groups of insects, whilst other important groups are neglected (Tautz et al., 2003). Further, certain groups of insects, particularly in their immature life stages, lack morphological characters with taxonomical information and diagnostic importance, and most juveniles, early instar and pupae cannot be identified (Martoni et al., 2021). Phenotypic plasticity, cryptic species, microbial species, metamorphosis, sexual dimorphism, camouflage and mimicry have complicated

morphological identifications and have resulted in a lot of confusions and downfalls (Friedheim, 2016; Tahir et al., 2018). Thus, molecular techniques have been used over the last decade as complementary and alternative methods to solve taxonomic problems in insect identification.

## Molecular identification of insects and its drawbacks

In recent years, molecular methods have been used to identify insects that are mainly important as pests and parasitoids of agriculture (Rasool et al., 2018; Jenkins et al., 2012), disease vectors of animals (Bakhoum et al., 2018) and humans (Collins et al., 2000), insect species associated with corpses which are important in forensic entomology (Gemmellaro et al., 2019), insect taxa required for biodiversity and conservation planning (Gomez-Zurita, 2016). Molecular methods that utilize protein and DNA markers have been widely used by taxonomists, ecologists, agriculturists and conservation biologists to discriminate insect species. The ability to use small, damaged, or even industrially processed insect material (Bakhoum et al., 2018; Rasool et al., 2018); the ability to gain information about the taxonomic affinities in the least morphologically amenable groups (Rasool et al., 2018); the ability to bypass the need for the specialist entomological knowledge required for morphological species differentiation (Jenkins et al., 2012); the ability to overcome difficulties in differentiation of morphologically similar species, larval forms or members of cryptic species complexes (Jenkins et al., 2012); less consumption of time (Bakhoum et al., 2018); and the provision of easier identification of host associations (Gaskin et al., 2011), have expanded the use of molecular techniques in insect identification. However, molecular identification data rely on good taxonomic, morphological, ecological and historical

information of insect taxa to be identified (Gaskin et al., 2011), and molecular tools may be expensive and require specialized equipment (Bakhoum et al., 2018). Further, the identification of a single species per reaction nature of most of these methods does not facilitate multi-species identification of complex mixed insect communities, allowing unanticipated species to avoid detection (Piper et al., 2019).

Given the drawbacks of morphological and molecular identification of insects and yet the need for easy and accurate identification in addressing concerns related to human food and health, there has long been an interest in developing machine learning systems for insect identification. In this approach, entomological observations such as phenotypic traits of insects, behavioural characters, and other interactions are connected with computer technology to identify different insect species and estimate their diversity, abundance and other factors.

## Machine learning systems for insect identification

Use of machine learning for automated species identification is an application of general pattern recognition, and it is a part of computer-aided taxonomy (CAT) (Chesmore, 1997). According to level of automation, automated species identification can be mainly divided into two levels:

1. Full automation - Complete identification without user interaction
2. Semi-automation - Automated identification using manual intervention by a domain expert

Semi-automation is more realistic than full automation as it allows prior sorting into higher taxonomic categories such as genera and more likely to be feasible in the short term. Further, semi-automated systems also allow user interaction up to some extent; for example, to select object location

within an image for identification. In terms of graphical design, an automated species identification system can be considered as a standard pattern recognition system, as shown in Figure 1, and it can be divided into four main functional blocks.

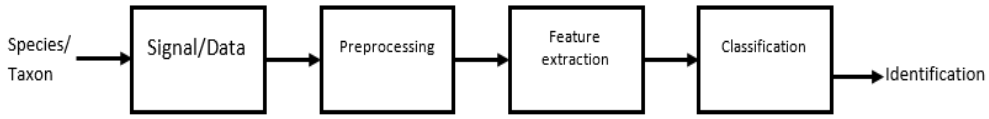



**Figure 1.** The four main functional blocks of an automated species identification system

## Signal/ data

There are a wide range of sensors to detect or measure different types of real-world signals/data, such as environmental parameters including temperature and humidity, chemicals, sound waves in land and water, radio waves and vibrations, etc. However, when it comes to automated species identification, only a limited number of signals can be used, since some signals are unable to provide useful data which can be utilized to distinguish between different species varieties. Some of the most common signals currently employed in modern species identification processes are acoustic signals that rely on the unique sounds of insects, E-noses/olfactory signals based on volatiles released from their bodies, and behavioural signals that depend on the movements of insects and visual signals.

### *Acoustic signals*

Many animals produce acoustic signals as a mean of communication or as a by-product of an activity such as flying, eating and locomotion. Most of

these acoustic signals occur within the human auditory frequency range (20 Hz–20 kHz), but some occur at lower frequencies known as infrasound (African elephants *Loxondonta africana* and various species of whale) (McComb et al., 2003), and higher frequencies termed ultrasound (20–20 kHz or higher: eg. bush-crickets, moths, grasshoppers). Sounds emitted by insects when feeding, moving or communicating have been used for developing potential applications for insect detection and identification in numerous occasions (Mankin et al., 2011). Many of these signals including the calling songs of cicadas and crickets, wing beat sounds of mosquitoes, tapping sounds of male beetles when communicating with the females, and communication sounds in different species of Hemiptera and Orthoptera have provided important information in identifying the insect species and its sex (Mankin et al., 2011).

Thus, the use of acoustic signals for insect identification expanded significantly during the past few decades and rapidly replaced many of the labor-intensive, less effective identification methods then in use. However, the lack of inexpensive, user-friendly acoustic tools for the identification of insect sounds, difficulties of interpreting weak insect signals in environments with high background noise, small market for insect acoustic detection devices, and the inability of signals to be detected over larger distances, depending on the substrate in which the insect is hidden has resulted in severe drawbacks of these methods in insect identification (Mankin et al., 2011). Acoustic sensing can be affected significantly by interference of background noise, mainly from other animals and anthropogenic sounds, where significant frequency overlap exists. Some automated acoustic signal-based species identification systems overcome this problem to an extent by distinguishing between anthropogenic sounds and those of the target taxa and filter the signal in situations where the

interfering sound has a much higher intensity that differs significantly from the target species (Chesmore and Ohya, 2004). Further, recent developments of acoustic technology incorporating neural networks and machine learning have enabled automated monitoring and distinguishing of stored product insects with success (Mankin et al., 2021).

### *Behavioural signals*

The first automated identification systems for insects were based on their movements and were developed as early as 1973. They used wingbeat frequency of flying insects, such as mosquitoes and flies of order Diptera, cicadas (Homoptera) and grasshoppers (Orthoptera), to identify species and their sex. At present, unsupervised automated sensors have been introduced to record insect flight movements and compute data that can be used to classify insects to the species level (Rydhmer et al., 2022). Flying insects such as sawflies and pod midges that infest oilseed rape crops have been classified with 80% accuracy using sensors combined with machine learning (Kirkeby et al., 2021).

### *Visual signals*

Automated image-based systems using visual signals have been developed since 1995 for insect identification (Valan et al., 2019). In these systems, relevant features of a visual image such as the wing veination pattern, the relative position of wing vein junctions, the outline of the wing or the whole body are extracted from the raw images by a system designed for feature extraction. Initially these systems were handcrafted and depended on manual procedures that were laborious and complex. In recent years, deep learning (DL) and convolutional neural networks (CNNs) have emerged

that can learn to extract relevant features, automatically, without human intervention (Valan et al., 2019). In these systems, a visual image is considered as a signal which contains a two-dimensional finite set of digital values, called picture elements or pixels (smallest unit of a digital image) (Figure 2), and are categorized into two types as black and white images (binary images) and colored images.

At present, the significant increase in the size and complexity of neural networks and the extra computational power brought by modern graphical processing units (GPUs) have generated spectacular advances in using visual signals in insect identification.

## Preprocessing

Preprocessing is the second step in an automated species recognition system and is the process of formatting different signal types before they are used by models for training and inference. The techniques of preprocessing which are necessary, are entirely dependent on the type, quality and quantity of the data/signal which is used for the species identification process. For example, some of the preprocessing techniques which are required for acoustic signals include:

Pre-emphasis filter - Amplifies high frequency sounds

Framing - Split audio signal into short-term windows called frames

Hamming - Reduce spectral leakage or any signal discontinuities and improve the clarity of the signal

Transforming - Conversion of continuous analog audio signal to a discontinuous digital form.

Some of the preprocessing techniques required for visual signals include:

Resize image - Images captured by different devices (mobile phone captures *versus* DSLR camera captures) can vary in size. Therefore, dimensions of the images are altered to one particular size.

Remove noise - Removal of unnecessary surrounding data (background details).

Segmentation - Split images into multiple parts or regions which are often based on the characteristics of the pixels in the image. Image segmentation could involve separating foreground from background, or clustering regions of pixels based on similarities in colour or shape.

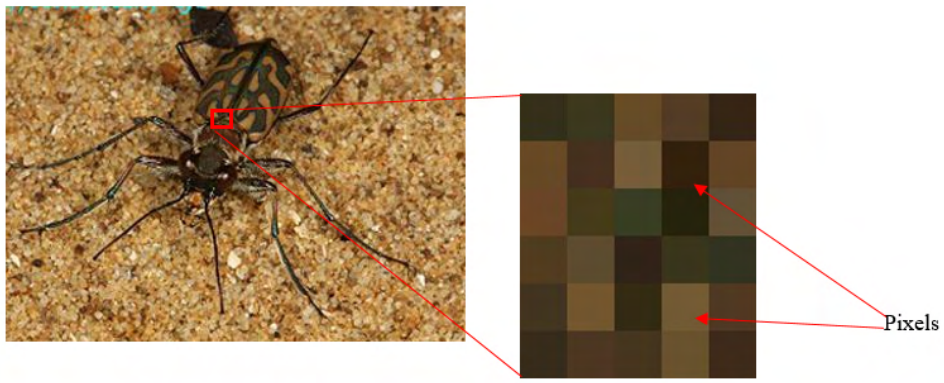



**Figure 2.** The pixels of a visual image


**Figure 3.** Feature extraction using visual signals

## Feature extraction

Feature extraction is the method that selects and /or combine variables in the preprocessed data/signals into features (individual measurable properties or characteristics), which reduce the amount of data that must be processed, while accurately and completely describing the original data set. The aim of feature extraction is to find the most compressed and informative set of features (distinct patterns) which enhance efficiency of the classifier. Furthermore, feature extraction is used to extract features from the original signal to accomplish reliable classification (Figure 3). Feature extraction is the most crucial part of signal classification because the classification performance can be degraded if proper features are not selected (Li et al., 2016; Subasi, 2019; Abeywardhana et al., 2020). This feature selection is also mentioned as dimension reduction, where a minimum subset of features is chosen from original set of features, which achieves maximum generalization ability. In order to accomplish this, several types of methods can be used, such as principal component analysis (PCA), independent component analysis (ICA), linear discriminant analysis (LDA), statistical values, and different entropy measures.

## Classification

The term 'classification', as used in computer science is defined as the assignment of a signal or a pattern to one of a number of pre-specified classes based on features extracted from the signal/pattern. Further, ability to classify these signals/patterns is known as ‘Recognition’ (Schalkoff, 1992). According to Schalkoff (1992) pattern recognition systems can be divided in to three main types. They are,

I. Statistical pattern recognition - A method which uses decision-theoretic concepts to distinguish objects belonging to different groups based upon their quantitative features.

II. Structural pattern recognition - A method which uses syntactic grammars to identify objects belonging to different groups based upon the arrangement of their morphological (shape-based or structural) features. Structural pattern recognition systems are difficult to apply into new domains because implementation of both the description and classification tasks requires domain knowledge.

III. Neural pattern recognition – A pattern recognition method developed in artificial intelligence inspired by biological neural network of the human brain. Currently, it is the most popular method for pattern detection. Artificial neural networks are based on parallel subunits referred to as neurons. It can be viewed as massively parallel computing systems consisting of a huge number of simple processors with many interconnections (neurons). Convolutional Neural Networks model which is one of the main models used in image-based species identification, is also referred under neural pattern recognition methods.

The main objective of automated systems is to achieve 100% accuracy for (unseen) data/signal which are going to be tested on top of the trained model. The trained model comprises the analysis of images or data that have been independently and accurately identified as taxa, and are now used to determine a classifier's parameters for providing maximum discrimination between taxa (Waldchen et al., 2018). However, in reality, this is impossible to achieve, especially when dealing with 'real' signals/data, which contain limited number of data (due to endemicity,

rareness) and background noise with high levels of interference etc. Therefore, the main goal of automated species identification is to achieve the highest most possible validation accuracy. In order to achieve that, it is important to identify exact potential limitations and design and implement a robust system with repeatable data to improve reliability of the model.

## Automated insect identification approaches in Sri Lanka

Insect identification using machine learning approaches have been conducted to a lesser extent in Sri Lanka. Most of the studies that have been conducted have concentrated on insect pest species of agricultural crops that are of economic importance to the country. Venugoban and Ramanan (2014), have developed an image-based machine learning approach to identify 20 insect pests of paddy in Sri Lanka, and the multi-class classification has been based upon support vector machine technique. A similar study also focusing on the paddy insect pests of the country has been carried out by Fernando (2021), in which images of insect species were classified using convolutional neural network models. However, a recent study considering an insect family with biodiversity value to the island and with a high number of endemic species has been taxonomically revised using machine learning approaches and has achieved immense success in recognizing unknown species (Abeywardhana et al., 2021; Abeywardhana et al., 2022). A large number of tiger beetle (family Cicindelidae) images were used for the study and features were extracted based on the different colours, shapes and textures of the species (Figure 4). Classical machine learning classifiers and deep learning techniques were used to classify tiger beetles into genera and species, and deep learning methods provided more accuracy.

Automated species identification systems have the potential to support taxonomists, entomologists, para-taxonomists in routine identification, naming, sorting, rapid biodiversity assessment and long-term conservation/ecological monitoring (Riede, 1993). Modern automated models would empower entomologists, nature enthusiasts to identify different species encountered during their field work, and employers like farmers to recognize unfamiliar insects in their fields. Insect species identification based on machine learning will provide a platform for the public to use taxonomic information and identification techniques that will empower them with the required knowledge. It will help to expand the knowledge in diversity of different species throughout the world and the same approaches will be an interface to face challenges of rapidly changing ecosystems, diversity and distribution of species.

**Figure 4.** Tiger beetle images used for automated identification (Abeywardhana et al., 2021)

## Drawbacks and challenges of automated species identifications

Despite the fact that machine learning techniques have been very effective in identifying insects, there are some common significant issues associated with reliable automated species identification.

One of the major difficulties in reliable automated species identification has been in obtaining a reasonably large and high-quality training set for species discrimination. Machine learning techniques using visual signals for insect identification tend to be affected on the background and lighting of images and pose of specimens in distinguishing species, and significant variations or too great a standardization are likely to result in misidentifications (Gaston and O'Neill, 2004). More research is needed to deal with noisy images, complex backgrounds, damage detection and digital image repair (Mata-Montero and Carranza-Rojas, 2016). Establishing larger training sets can be difficult where some of the species to be included are rare, and specimens may be difficult to obtain (Gaston and O'Neill, 2004).

Further, an automated taxon identification approach not only needs to be able to match an individual specimen to one of the known taxa, but should also be able to reject specimens belonging to a taxon that was not part of the training set. However, finding a machine learning approach with such characteristics have also posed challenges in classifier design and training (Waldchen et al., 2018).

A computational problem is also seen in insect discrimination, as automated identification is computationally intensive, and as the number of possible species to be identified increases, the acquisition of sufficient computing power becomes limited (Gaston and O'Neill, 2004).

Machine learning techniques are known to give errors in identification for insect groups that have poor taxonomic information or which are distinguished using single morphological structures (Gaston and O'Neill, 2004).

Automated species identifications do not entirely remove the need for human involvement, and material for identification needs to be prepared, images obtained, data preprocessed, and features extracted. Therefore, in establishing identifications of many specimens, substantial amounts of work are involved making the process labour intensive (Gaston and O'Neill, 2004).

Further, fully functional automated identification systems are complex combinations of hardware and software, and the implementation of a production system is an expensive undertaking (Gaston and O'Neill, 2004). However, when considering the opportunities that machine learning will provide for species recognition, it would be valuable to overcome these obstacles.

## References

Abeywardhana, DL, Dangalle, CD, Nugaliyadde, A & Mallawarachchi, Y 2022. An ultra-specific image dataset for automated insect identification. *Multimedia Tools and Applications*, 81(3), pp.3223-3251. https://doi.org/10.1007/s11042-021-11693-3

Abeywardhana, DL, Dangalle, CD, Nugaliyadde, A & Mallawarachchi, Y 2021. Deep learning approach to classify tiger beetles of Sri Lanka. *Ecological Informatics*, 62. https://doi.org/10.1016/j.ecoinf.2021.101286

Abeywardhana, DL, Dangalle, CD, Nugaliyadde, A & Mallawarachchi, YW 2020. Analysing feature importance for identification of tiger beetles using machine learning. *Proceedings of the Annual Research Symposium of the University of Colombo*, pp. 43.

Bakhoum, MT, Sarr, M, Fall, AG, Huber, K, Fall, M, Sembene, M, Seck, MT, Labuschagne, K, Gardes, L, Ciss, M, Gimmoneau, G, Bouyer, J, Baldet, T &

Garros, C 2018. DNA barcoding and molecular identification of field-collected *Culicoides* larvae in the Niayes area of Senegal. *Parasites & Vectors*, 11. https://doi.org/10.1186/s13071-018-3176-y

Chan, A, Chiang, LP, Hapuarachchi, HC, Tan, CH, Pang, SC, Lee, R, Lee, KS, Ng, LC & Lam-Phua, SG 2014. DNA barcoding: complementing morphological identification of mosquito species in Singapore. *Parasites & Vectors*, 7. http://www.parasitesandvectors.com/content/7/1/569

Chesmore, ED 1997. Methodologies for automating the identification of species. In: Gallagher S ed. *Proceedings of the inaugural meeting of the BioNET-International group for computer-aided taxonomy (Bigcat), 2-3 July 1997, The University of Wales, Cardiff.* Surrey: BioNET-International Technical Secretariat, pp.3-12.

Chesmore, ED & Ohya, E 2004. Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. *Bulletin of Entomological Research,* 94(4), pp.319-330. https://doi.org/10.1079/BER2004306

Collins, FH, Kamau, L, Ranson, HA & Vulule, JM 2000. Molecular entomology and prospects for malaria control. *Bulletin of the World Health Organization*, 78, pp.1412-1423.

Eidt, DL 1995. The importance of insect taxonomy and biosystematics to forestry. *The Forestry Chronicle,* 71, pp.581-583.

Fernando, FFSN 2021. Paddy pest recognition based on deep learning in Sri Lanka. A dissertation submitted for the Degree of Master of Computer Science, University of Colombo School of Computing.

Friedheim, S 2016. Comparison of species identification methods: DNA barcoding versus morphological taxonomy. *Manoa Horizons,* 1, pp.74-86. https://kahualike.manoa.hawaii.edu/horizons

Gaskin, JF, Bon, MC, Cock, MJW, Cristofaro, M, De Biase, A, De Clerck-Floate, R, Ellison, CA, Hinz, HL, Hufbauer, RA, Julien, MH & Sforza, R 2011. Applying molecular-based approaches to classical biological control of weeds. *Biological Control,* 58, pp.1-21. http://doi:10.1016/j.biocontrol.2011.03.015

Gaston, KJ & O'Neill, MA 2004. Automated species identification: why not?. *Philosophical Transactions of the Royal Society B,* 359, pp.655-667. https://doi.10.1098/rstb.2003.1442

Gemmellaro, MD, Hamilton, GC & Ware, JL 2019. Review of molecular identification techniques for forensically important Diptera. *Journal of Medical Entomology*, pp.1-16. https://doi:10.1093/jme/tjz040

Gomez-Zurita, J, Cardoso, A, Coronado, I, De la Cadena, G, Jurado-Rivera, JA, Maes, JM, Montelongo, T, Nguyen, DT & Papadopoulou, A 2016. High-throughput biodiversity analysis: Rapid assessment of species richness and ecological interactions of Chrysomelidae (Coleoptera) in the tropics. *ZooKeys*, 597, pp.3-26. https://doi:10.3897/zookeys.597.7065

Jenkins, C, Chapman, TA, Micallef, JL & Reynolds, OL 2012. Molecular techniques for the detection and differentiation of host and parasitoid species and the implications for fruit fly management. *Insects*, 3, pp.763-788. https://doi:10.3390/insects3030763

Kirkeby, C, Rydhmer, K, Cook, SM, Strand, A, Torrance, MT, Swain, JL, Prangsma, J, Johnen, A, Jensen, M, Brydegaard, M & Graesboll, K 2021. Advances in automatic identification of flying insects using optical sensors and machine learning. *Scientific Reports*, 11, 1555. https://doi.org/10.1038/s41598-021-81005-0

Li, H, Liang, H, Miao, C, Cao, L, Feng, X, Tang, C & Li, E 2016. Novel ECG signal classification based on KICA nonlinear feature extraction. *Circuits, Systems, and Signal Processing*, 35 (4), 1187–1197. http://dx.doi.org/10.1007/s00034-015-0108-3

Mankin, RW, Hagstrum, DW, Smith, MT, Roda, AL & Kairo, MTK 2011. Perspective and promise: a century of insect acoustic detection and monitoring. *American Entomologist*, 57, pp.30-44.

Mankin, R, Hagstrum, D, Guo, M, Eliopoulos, P & Njoroge, A 2021. Automated applications of acoustics for stored product insect detection, monitoring, and management. *Insects*, 12, 259. http://doi.org/10.3390/insects12030259

Martoni, F, Valenzuela, I & Blacket, MJ 2021. On the complementarity of DNA barcoding and morphology to distinguish benign endemic insects from possible pests: the case of *Dirioxa pornia* and the tribe Acanthonevrini (Diptera: Tephritidae: Phytalmiinae) in Australia. *Insect Science*, 28, pp.261-270. https://doi.10.1111/1744-7917.12769

Mata-Montero, E & Carranza-Rojas, J 2016. 'Automated plant species identification: Challenges and opportunities', in *IFIP International Federation for Information Processing: proceedings of the IFIP ACIT conference*, eds. F.J. Mata & A. Pont, WITFOR, Switzerland, 481, pp.26-36. http://doi.10.1007/978-3-319-44447-53

McComb, K, Reby, D, Baker, L, Moss, C & Sayialel, S 2003. Long-distance communication of acoustic cues to social identity in African elephants. *Animal Behaviour,* 65, pp.317-329. https://doi:10.1006/anbe.2003.2047

Piper, AM, Batovska, J, Cogan, NOI, Weiss, J, Cunningham, JP, Rodoni, BC & Blacket, MJ 2019. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *GigaScience,* 8, pp.1-22. https://doi:10.1093/gigascience/giz092

Rasool, A, Ahmad, T, Ganai, BA & Gull, S 2018. An overview of molecular identification of insect fauna with special emphasis on chalcid wasps (Hymenoptera: Chalcidoidea) of India. *Acta Agriculturae Slovenica,* pp.229-239. https://doi:10.14720/aas.2018.111.1.22

Rydhmer, K, Bick, E, Still, L, Strand, A, Luciano, R, Helmreich, S, Beck, BD, Gronne, C, Malmros, L, Poulsen, K, Elbaek, F, Brydegaard, M, Lemmich, J & Nikolajsen, T 2022. Automating insect monitoring using unsupervised near-infrared sensors. *Scientific Reports,* 12, 2603. https://doi.org/10.1038/s41598-022-06439-6

Subasi, A 2019. *Practical guide for biomedical signals analysis using machine learning techniques: A MATLAB based approach*. Academic Press.

Tahir, HM, Noor, A, Mehmood, S, Sherawat, SM & Qazi, MA 2018. Evaluating the accuracy of morphological identification of insect pests of rice crops using DNA barcoding. *Mitochondrial DNA Part B,* 3, pp.1220-1224. https://doi.org/10.1080/23802359.2018.1532334

Tautz, D, Arctander, P, Minelli, A, Thomas, RH & Vogler, AP 2003. A plea for DNA taxonomy. *TRENDS in Ecology and Evolution,* 18, pp.70-74.

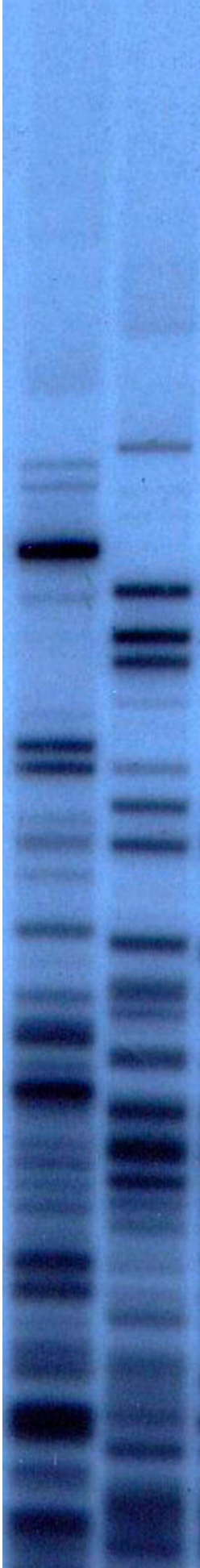
Valan, M, Makonyi, K, Maki, A, Vondracek, D & Ronquist, F 2019. Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Systematic Biology,* 68, pp.876-895. DOI:10.1093/sysbio/syz014.

Venugoban, K & Ramanan, A 2014. Image classification of paddy field insect pests using gradient-based features. *International Journal of Machine Learning and Computing,* 4(1), pp.1-5. https://doi.org/10.7763/IJMLC.2014.V4.3761

Waldchen, J, Rzanny, M, Seeland, M & Mader, P 2018. Automated plant species identification -Trends and future directions. *PLoS Computational Biology,* 14 (4). https://doi.org/10.1371/journal.pcbi.1005993

# Chapter 2

# Changing paradigms in forensic genetics



K. L. N. Perera[1]*, G. Galhena[2]

[1]Genetech Molecular Diagnostics, Colombo 08, Sri Lanka
[2]Department of Zoology and Environmental Science, University of Colombo, Colombo 03, Sri Lanka
* Corresponding author:
nandika@genetechsrilanka.com

## Abstract

The introduction of DNA fingerprinting technology signifies the first shift in the changing paradigms of forensic science. The method revolutionised the realm of person identification by tracking individuals with unbelievable accuracy, using only traces of biological material. Throughout the last four decades, DNA fingerprinting has made an enormous contribution to the judiciary, facilitating the conviction of criminals, exoneration of the wrongly accused, identification of victims of mass disasters and resolution of kinship disputes. This chapter focuses on the changing landscapes of forensic genetics over the years, expanding its scope from the classical STR profile comparisons to tracing perpetrators, who did not leave behind a trace. The transformations brought in by incorporating cutting-edge genetic technologies to forensic ecology, wildlife forensics, and food adulteration have also been discussed together with their applications and challenges. Integration of the forthcoming RNA revolution in addressing the crucial aspects of forensic scenarios has also been explored.

## Keywords

*DNA Fingerprinting, Environmental DNA, Kinship, Microsatellites, Next Generation Sequencing*

## A transcending science

Forensic Science is the realm where scientific method is applied in the judicial system to resolve disputes and enforce law in both civil and criminal proceedings. It draws upon a variety of scientific disciplines including biology, physics and chemistry to different extents. The renaissance of forensic sciences was marked by the exchange principle put forwarded by the French criminologist Edmond Locard in early 20$^{th}$ century which states that “every contact leaves a trace”. The principle is interwoven with the fundamental question raised in forensic science - the identity of individuals. In criminal cases like murder, sexual assault or burglary, there are offenders that need to be identified, while in more sophisticated cases like disputed paternity, incest, child adoption and victim recognition in mass disasters, kinship between individuals need to be assessed. In all these instances, biological materials are typed using various methods to establish the individual identity. The conventional methods used for the typing of biological material such as examination of dental records (odontological evidence) and typing of tissue and body fluid using various biochemical markers have a very narrow use in these instances, while fingerprints (chiral) although can help in criminal investigations, its availability can easily be avoided. In this context, the use of DNA evidence for person identification has been the single most important event that took place within the last century in the field of forensics. The practice turned a new leaf, and the method was taken up immediately by the forensic investigators all over the world marking the birth of a brand-new field, “forensic genetics”.

## Paving the way for forensic genetics

Although DNA evidence was first used in forensic casework as recently as in mid 1980s, the first tool which distinguished between individuals based on their genetic characters was introduced more than a century ago. This tool was the ABO (formerly ABC) blood grouping system. Although by 1902 all four main blood groups (A, B, AB and O) had been identified, the testing could not be integrated to forensics until Leone Lattes, a serologist in Italy introduced a method to determine the blood type in 1915. Even though this system produced a convenient and fast method of categorizing people, it was not very informative as a method of person identification, other than when excluding an individual from being the source of a crime scene sample. This is due to its inability to provide a unique identification signature for a given individual. However, it did help in tracking family relationships up to some extent.

After the establishment of variation in blood groups, a new era of haemogenetics was established where individuals were characterized for their genetic constituents based on biochemical markers. This was facilitated by the rapid development of electrophoresis techniques such as zone electrophoresis, immune-electrophoresis and isoelectric focusing which were able to separate proteins based on their numerous characteristics. With the development of these techniques many types of polymorphisms such as polymorphism in haptoglobin (HP), serum proteins polymorphisms, erythrocyte enzyme polymorphisms, and the polymorphism in Human Leucocyte Antigen (HLA) system started to unravel. Even though these polymorphic systems improved the discrimination power of the forensic analysis, it was not sufficient for individualization. In addition, the proteins were not present at sufficient levels to carryout typing in most tissues collected as forensic samples.

Further, they tend to be relatively unstable outside the body, restricting their typing in crime scene samples. Although typing genetic polymorphisms at the level of DNA could help overcoming many of these limitations, the possibility was opened up only after Watson and Crick proposed their Nobel prize winning model for the genetic blueprint in 1953. Since then, the human genome started to unravel at an astounding speed and in parallel the field of forensic genetics started revolutionizing as never before.

## Introduction of DNA fingerprinting

In the 1980s, extensive research conducted on the human genome led to the discovery of large repetitive DNA elements which can vary in size from 10 - 100 bp, in the regions outside the genes of nuclear DNA. These hypervariable minisatellite regions which are also known as VNTR (variable number of tandem repeats) was the basis for first forensic DNA analysis published by Dr. Alec Jeffreys in 1985. The analysis of VNTR was conducted with the restriction fragment length polymorphism (RFLP) technique, which therefore had a cumbersome procedure. Jeffreys called it DNA “fingerprinting”.

The first application of VNTR based DNA fingerprinting was an immigration case which saved a young United Kingdom (UK) born Ghanaian boy from deportation in 1985 (Jeffreys et al., 1985). The reputation paved way for the technique to be applied in forensic investigations for the first time in 1986 to resolve a rape case of two teenage girls that took place in a village at Leicestershire, England. Typing VNTR loci by RFLP analysis provided a very high power of discrimination. Nevertheless, the method required samples containing at least 10–25 ng of relatively intact DNA with fragment lengths up to 10,000 bp for successful

typing. Forensic samples often consist of trace amounts of degraded DNA and may not comply this requirement. In addition, VNTR profiles had numerous technical problems when used in routine analysis (Lewontin and Hartl, 1991). For example, alleles that differed by only a few repeats could not be separated reliably with electrophoresis and produced a quasi-continuous array. Although this was solved by defining 'bins', the bands assigned to the same bin contained different numbers of repeats. Further, the rate of migration of identical DNA fragments varied from gel to gel, making it difficult to compare among gels. All these shortcomings of VNTR/RFLP method pointed at the requirement of a more stringent procedure for accurate forensic analysis.

## The impact of polymerase chain reaction (PCR)

The discovery of PCR amplification of DNA in 1983 had a huge impact on all molecular methods and revolutionised the scope of DNA analysis. Being more sensitive, accurate and fast, PCR based methods soon overpowered the other methods which made use of genomic DNA directly.

The first human genetic marker system analyzed using PCR-based technique involved the detection of single nucleotide polymorphism (SNP) using synthetic allele specific oligonucleotide (ASO) probes (Saiki et al., 1986). In this method, sequence polymorphisms at the HLA-DQA1 locus and polymarker loci (LDLR, GYPA, HBGG, D7S8, and Gc) were typed using ASO probes in a reverse dot blot format. The method incorporated a PCR amplification step before hybridization which significantly increased the sensitivity of the test. Even though this SNP-based system offered a higher sensitivity, it had a lower power of discrimination than the VNTR/RFLP typing method. This led to the attempts of amplifying VNTR

loci using PCR to produce a more advanced system that could incorporate both the sensitivity of PCR and the discrimination power of VNTRs.

## STR analysis: a new paradigm

Microsatellite analysis was introduced in this background and paved way to a new shift in forensic genetics. Microsatellites (also known as short tandem repeats: STRs) are a subclass of VNTR loci, that consist of tandem repeats of two to five bases long DNA motifs. They are distributed throughout the genome roughly at every 10,000 nucleotides and constitute approximately 3% of the total human genome. Thus, microsatellites were considered a potentially promising tool that could capture the genetic variability among individuals.

STR typing was introduced as a PCR based technique and had many advantages over the first-generation DNA typing methods, i.e. VNTR/RFLP method and PCR based HLA DQA typing method (Roewer, 2013). STR typing encompassed all the advantages that PCR technique could impart such as higher accuracy, sensitivity, specificity and speed. In addition, the method was comparatively more convenient and was less prone to allelic dropout than VNTR systems. Since DNA fragments generated using STR markers are relatively short, they are more successful in investigating degraded DNA which is an added advantage of the method. Thus, soon after its first introduction in early 1990s, STR typing became the method of choice for human identification.

To use STR typing as the routine method of DNA fingerprinting, it was necessary to select a core set of STR loci for the use of forensic analysis. Such a consensus set of loci will ensure the effectiveness of DNA profiling across a wide array of jurisdictions and can also facilitate the construction

of national and international databases to allow comparison of DNA profiles worldwide. Identifying this need, FBI (Federal Bureau of Investigation, USA) laboratory sponsored a community-wide forensic science effort in 1996, to establish a core set of STR loci for inclusion within the national DNA database. The effort identified 13 core autosomal STR loci which came to be known as the CODIS (Combined DNA Index System) core loci (Butler, 2011). Parallel to this development, European forensic DNA laboratories also came up with a standard set of loci known as the European Standard Set (ESS) which shared seven autosomal STR loci in the CODIS loci. Later on, both the standards were extended to include more loci, i.e. CODIS now has 20 core loci and ESS-extended has 12 loci.

To facilitate the co-amplification of this increased number of STR markers included in the standard set, single tube multiplex PCR systems were soon developed and were made available commercially. With the introduction of automated capillary gel electrophoresis systems coupled with fluorescence detection, a single capillary gel became capable of producing a full STR profile of an individual, making the process of individual identification a relatively fast and a convenient one.

## Forensic genetic databases

Once there was consensus among the forensic community on a standard set of autosomal STR markers, it became possible to construct forensic DNA databases for the use of law enforcement systems. Usually, these forensic DNA databases store DNA profiles of criminal offenders deemed to legally qualify to enter the database. During criminal investigations, the databases are searched to look for a potential match for a crime scene sample. The effectiveness of a database lies on the fact that majority of crimes are

committed by repeat offenders and its performance becomes better when the database increases in its size.

The establishment of an effective DNA database requires a collaborative effort from several parties including forensic DNA laboratories, the law enforcement community, and government policy makers. The first national forensic DNA database in the world was established in the United Kingdom in 1995. Its tremendous success in resolving crimes convinced many other countries to initiate their individual national forensic databases. At present, about 60 countries in the world have established forensic DNA databases with another 34 countries including several South Asian countries like India, Thailand and Pakistan planning to set up new DNA databases. China (20 million profiles) and the United States (12 million profiles) maintain the two largest forensic DNA databases with a production over 410 000 and 185 000 hits respectively (Ge et al., 2014). In addition, there are international DNA databases which allow the exchange of DNA profiles between countries. For example, the INTERPOL DNA database established in 2002 has more than 247,000 DNA profiles contributed by 84 member countries participating in it, as per its website. Likewise, there are regional DNA databases for selected communities such as Europol Information System (EIS) which includes profiles from European Union (EU) Member States. This growing number of DNA databases around the world has largely revolutionized the ability to link crime scene evidence to known perpetrators.

## Statistical validity of DNA base human identification

Parallel to the technological advances that took place in forensic genetics, there was a major development in the statistical approaches used in the

analysis. Instead of the potentially misleading match/no match claims practiced in early days of forensics science, the later development attempted using more plausible probabilistic calculations. This approach tries to identify the statistical probability at which the evidence of interest that matched with the suspect could actually have originated from the suspect. This is achieved via calculating what is generally known as the "match probability", which is the probability that a random person in the population will have the same profile. During the first phase of DNA fingerprinting which involved VNTR/RFLP analysis the probative value of a match is often calculated by multiplying together the estimated frequencies with which each particular VNTR pattern occurs in a reference database. However, this method was liable to potentially serious errors because 1) ethnic subgroups within a country or a race exhibit genetic differences that are maintained by intermarriages within smaller subgroups and 2) due to the linkage disequilibrium that could exist among markers. In 1991, Lewontin and Hartl proposed a set of revisions to correct these drawbacks which was later incorporated in to the STR typing analysis as well. With these improvements, the combination of all 13 CODIS core loci produces an average random match probability value lower than one in trillion among unrelated individuals (Chakraborty et al., 1999). This probability far exceeds the population of earth making it virtually impossible for two individuals who are not genetically identical to have the same CODIS profile and is considered the greatest strength of DNA typing.

## Markers on sex chromosomes

The next milestone of DNA fingerprinting application was marked by the integration of gonosomal (X and Y chromosomes) STR repertoire to the

routine analysis. Although CODIS markers offered a highly efficient analysis system for crime scene investigations, they have lapses in providing conclusive evidence related to kinship analysis at specific instances. The inheritance patterns of the X and the Y chromosomes differ from the autosomes, as well as from each other. Consequently, the markers located on them carry the potential to offer different but complementary information necessary for both forensic and kinship applications. For instance, Y-STR typing has certain advantages over the autosomal STRs in obtaining a male profile in the presence of female DNA in cases like sexual assaults. This is possible since Y-STR haplotype can be specifically targeted without the interfering effect from female DNA. Y-STR typing is also useful over autosomal typing in deficient paternity cases of male offspring by facilitating the use of paternal male relatives in proving the paternity. On the other hand, X chromosomal STR (X-STR) markers complement the analysis of complex kinship scenarios, where at least one female is involved. These scenarios commonly include, but not limited to, deficient paternity cases of female children, cases with female siblings sharing a common biological father, relationship disputes concerning paternal grandmother-granddaughter and other distant female relatives. Moreover, X-STR analysis can serve in forensic case work in which female traces are to be identified in male background contamination.

## Contribution from mitochondrial DNA (mtDNA) analysis

Even with the advances that took place in the field of STR analysis, those samples that contain too little template DNA or are extremely degraded are not amenable to the analysis. In such instances, analysis of mtDNA can be of help due to its superior sensitivity. On average, each cell contains

between $10^3$ and $10^4$ copies of the mitochondrial genome, in contrast to the 23 pairs of chromosomes targeted by the STR typing. However, the discrimination power of mitochondrial analysis is very much less than STR typing, therefore cannot provide a basis for person identification on its own. In kinship analysis however, it helps in identifying the maternal lineage, though the search cannot be narrowed down further to pick up the exact relationship within the maternal line. Further, mtDNA typing is laborious, time consuming, and costly compared to STR typing since it requires sequencing of selected regions with high individual variability. Usually, hypervariable regions I and II (HVI and HVII) are commonly used in mtDNA analysis while HVIII is also additionally sequenced, where same haplotypes are observed in both HVI and HVII (Lutz et al., 2000).

## The role of single nucleotide polymorphisms (SNPs)

SNPs outperform both STRs and mtDNA and are the marker choice in analysing highly degraded samples. SNPs allow designing assays with very small amplicon sizes of 60–80 bp length (Budowle et al., 2003) and thus are capable of capturing the variation in even severely degraded evidentiary samples. Occurring almost once in every 1,000 nucleotides on average, there are roughly 4 to 5 million SNPs in a person's genome making it a very rich source of individual variability. However, most SNPs are biallelic and therefore not very informative like STR markers which have many alleles. Because of this, many SNPs are needed to get the same power of discrimination provided by a single STR loci, i.e. to obtain the power of discrimination of a STR multiplex system with 13 markers, a panel of at least 50–100 SNP would be required. In addition, owing to the limited

number of alleles per locus, SNPs are more problematic for interpretation in situations involving mixed samples.

Nonetheless, with the recent introduction of microarray technology ("DNA chips") that enables typing massive number of SNPs in parallel, the utility of SNPs has expanded beyond the expectation of conventional forensic typing. A SNP microarray carries thousands to millions of nucleic acid fragments with known SNPs, immobilized on to a solid surface. When these DNA strands come in contact with DNA isolated from the sample of interest, complementary strands between the sample and the chip anneal with each other. Each such hybridization event will emit a fluorescence signal that allows its identification.

The power of SNPs in forensics lies in its potential to provide information related to variety of different scenarios. There are four categories of SNPs that are identified as suitable for different types of forensic investigations. These include SNPs which allow for individualization (identity testing SNPs), assessment of kinship (lineage informative SNPs), establishing biogeographical ancestry (ancestry informative SNPs) and predicting phenotypic characteristics (phenotype informative SNPs) (Budowle et al., 2003). Several of these SNP-based tests have already been commercialized. For example, there are several SNP microarray kits for eye colour, paternity (http://www.dnaprint.com) and ancestry. However, the amount of information accessible through some of these tests are relatively restricted, i.e., ancestry informative SNP arrays are only capable of quantifying mixtures of geographic ancestry from four groups, African, Indo-European, Native American and East Asian. Likewise, phenotyping is only possible for limited number of traits such as eye, hair and skin colour with a lower predictability for intermediate traits. Nevertheless, the usefulness of such

information cannot be underestimated in a forensic setting where relevant details of a criminal are not available.

Originally SNP microarrays were developed for relatively high input DNA quantity and quality such as the DNA obtainable when diagnosing of diseases. This feature restricts its application for forensic DNA which typically tends to be low in quantity and quality. However, more recent developments in SNP arrays that focused on obtaining successful SNP signatures from trace amounts of compromised DNA and DNA mixtures have greatly enhanced the utility of SNP arrays in forensic investigations (de Vries et al., 2022). Nevertheless, despite their expanding utility, SNPs are not expected to replace STRs in routine forensic genetic analysis in the near future due to the rarity of SNP databases available for legal enforcement activities and their relative high cost compared to capillary electrophoresis based STR analysis.

## Potential of next generation sequencing in forensics

Recent developments in next-generation sequencing (NGS) technologies have significantly improved the scope of nucleic acid analysis over the past decade. NGS is an umbrella term that includes second-generation sequencing technology based on loop array sequencing, which can analyze a large number of samples simultaneously (massively parallel sequencing: MPS) as well as third-generation sequencing technology, which can determine the base composition of single DNA molecules in real time (SMRT: Single molecule real time) (Yang et al., 2014). These technologies have paved way for the next paradigm shift in the forensic genetics by removing the limits imposed by conventional analytical methods.

Using NGS technology, millions or billions of DNA molecules can be sequenced in a short time, thereby increasing the throughput substantially compared to what is possible with conventional Sanger sequencing or fragment analysis used in STR genotyping. This ability to generate unprecedented amount of information has opened up a major new frontier for forensic science. NGS analyses numerous marker types including STR, SNP as well as mtDNA, all at once, covering multiple aspects of individual identification. In addition, large scale sequencing of specific RNA subtypes, such as long non-coding RNAs and snoRNA (small nucleolar RNA), as well as methylated DNA, allow for testing of epigenetic markers, which can potentially yield valuable information about the age and lifestyle of suspects. NGS further allows for mixtures of genomes of any species to be sequenced in one analysis (Ansorge, 2009) which otherwise requires several amplification rounds aimed specifically at each species. This feature had extensively been utilized in the present technological leap seen in the field of forensic ecology as explained later.

Recognizing these advantages, many countries have now validated NGS technique within their accredited workflows applicable for investigative agencies and crime labs. Accordingly, commercial kits have been developed for simultaneous analysis of STR markers, SNPs and mitochondrial analysis using NGS. For example, the ForenSeq DNA Signature Prep Kit (Verogen, CA) is the first commercial forensic kit for the MiSeq FGx (an NGS platform validated for forensic use) and it allows for amplification of up to 153 (DNA primer mix A) or 231 (DNA primer mix B) markers simultaneously. The markers consist of 27 forensic autosomal STRs, 24 Y-STRs, 7-X STRs, the Amelogenin sex marker, and either 94 or 172 SNPs depending on the multiplex formulation. Among these 172 SNPs, there are informative identity, geographical ancestry and

phenotypic SNPs. Likewise, Thermo Fisher Scientific, a USA based company, has also designed and released a variety of Precision ID Panels for use on the Ion Torrent platform (a type of NGS platform introduced by Thermo Fisher Scientific) including those focused on STRs, identity SNPs, ancestry SNPs, and mitochondrial sequencing. In addition to the availability of increasingly more NGS based forensic test kits and platforms, it is also of utmost importance to integrate them within the legislative framework. A major breakthrough with this regard is the recent approval granted by FBI for the uploading of DNA profiles generated by Verogen (forensic Biotechnology company based on California) forensic technology into the National DNA Index System (NDIS) DNA profile database (SWGDAM, 2019) making it the first NGS technology approved for NDIS.

The use of NGS technology has reduced the risk of a false identification to astronomically low levels. For instance, the identification power of a full 21-locus profile is exceptionally huge with a chance of a random match 1 in $10^{29}$ or $10^{30.}$ When comparing this probability with the world population, which is only about $8 \times 10^9$, the reliability of a NGS match is beyond one's expectation. The added advantage of having such a big discrimination power is that even when a crime scene sample yields only a partial profile, there will still be enough information to make a 'match' with a database sample or with a suspect.

The incorporation of NGS methods have also increased the sensitivity of the forensic testing protocols in an astounding manner. With NGS, template amounts of under 100 picograms (15 cells' worth of DNA) will be sufficient to generate a full profile in contrast to the 1 ng needed for the routine PCR based fingerprinting technologies. Impressively, all these advantages come with a much lower cost, although initial investment for instrument is

substantial. However, with the latest developments of the technology, the cost of the NGS machines is becoming increasingly lower and comparable to the automated capillary electrophoresis based sequencing machines, NGS technology is greatly increasing the cost-effectiveness of legal cases.
In addition to analysing human samples, NGS technology has revolutionised the use of environmental trace evidence in forensic analysis. This field which is often referred to as forensic ecology is very useful in establishing a link between a suspect, location and victim in criminal cases. Forensic ecological investigations characterise biological specimens such as plant fragments, pollens, diatoms, fungi or insects from crime scene samples and match them with samples obtained from suspect's belongings. In a traditional analysis, the match is based on physical characterisation of such specimens and requires a diverse range of expertise including botanists, palynologists, mycologists and entomologists. Due to this limitation posed by the number of experts necessary, forensic ecology has not been used often enough in resolving criminal cases, despite its massive potential.
With the introduction of NGS, many of these obstacles imposed by the classical analysis techniques were lifted and consequently NGS has now become a key technique in the field of forensic ecology. With NGS at its core, the focus of the analysis is shifted from morphological traits to DNA sequence-based characterisation of biological material. This has coined a new term "environmental DNA (eDNA)" which describes the trace DNA elements present in the environment. In addition to plants (including pollen and diatoms) and invertebrates that are described in classical forensic ecology, eDNA encompasses microbes as well and have expanded the boundaries of forensic ecology further.

The most common method applied for characterization of complex genetically diverse eDNA samples is termed ‘DNA metabarcoding’. It involves sequence-based detection of a taxonomically informative gene termed a ‘barcode’, using universal primers devised for a single taxonomic group. Since metabarcoding make use of MPS technique, characterisation of microbes, plants and invertebrates are all done simultaneously from a complex eDNA mixture without having to isolate individual specimens. Thus, the requirement for calling the services from all kinds of specialties to identify the constituent biological components of a forensic sample no longer exists.

Interestingly, NGS has enabled the detection of trace levels of DNA in the environment even from low biomass samples such as dust. When objects are exposed to a particular environment, trace amounts of eDNA can be deposited on their surfaces as dust. Conversely eDNA can also be transferred from an individual by touch. Likewise, when two items come into contact, they will exchange the eDNA on their surfaces. Although such eDNA is present at very low levels, rather different eDNA signatures have been observed in different geographical areas and also in different habitat locations like households, offices and subway networks (Young and Linacre, 2021). In addition, different airborne pollen signals were detected from different surfaces like ropes, twines, clothing, fabrics, drugs, air filters, plant material, etc. Likewise, individual microbiome of each person tends to be different with partial similarities reflecting habitat co-sharing. Therefore, eDNA analysis conducted through MPS is regarded a very sensitive and a highly discriminative technique for provenance assignment, post-mortem interval estimations (time since death), human identification, body tissue identification and sexual assaults. However, full integration of eDNA into forensic genetics needed robust protocols for standardization,

validation and results comparison. Likewise, the necessity of developing a robust probability-based likelihood estimation to move away from the match/no match-based analysis framework is also essential for forensic ecology as well. Although some attempts have been made in this regard during recent years, a consensus has not been reached so far. This signals the necessity for more efficient and advanced systems that could address all existing uncertainties to their full extent.

## The plight of no suspects

The biggest challenge in forensic genetic applications has been the identification of unknown perpetrators in the absence of a suspect. This happens when there is no plausible evidence to suspect an individual as the perpetrator to test their DNA profiles for a match or when DNA profiles from crime scene do not match to any known suspect. Such instances lead to cold cases that are difficult to resolve allowing perpetrators to continue their criminal activities over extensive periods.

To enhance the power to identify unknown perpetrators, many DNA-based approaches were put forward during the past decade, among which ‘familial search’ was the most prominent. In this method, STR-profile generated from the crime scene is used to search for family members of the perpetrator from forensic DNA database. However, this approach could not generally trace beyond the first-degree relatives such as parents-offspring or full siblings (with 50% DNA sharing) or at most, avuncular relationships such as uncle-nephew (25% shared) due to the limited number of autosomal STRs used in routine forensic DNA profiling. However, Y-chromosome STRs (Y-STRs) haplotypes can overcome this limitation and because a vast majority of criminals are males, it has a far-reaching utility over autosomes

in such an instance. Likewise mitochondrial DNA haplotype analysis can also help in tracing maternal relatives. Nevertheless, since routine forensic analysis consists only of autosomal STR profiles and do not include Y-STRs or mtDNA, database searches based on these markers are of little value.

To extend familial search to cover a longer range of pairwise comparisons that include first cousins, second cousins, third cousins and beyond (12.5%, 3.13% and 0.78% DNA shared, respectively), genetic variation at much higher densities than the standard forensic tests are required. High-density microarray genotyping which consist of more than 600,000 SNPs comes in handy in this situation and provides the necessary information to make matches with both close and distant relatives. Known as investigative genetic genealogy (IGG; also known as forensic genetic genealogy: FGG), this technique opened up a promising new avenue to pursue unknown perpetrators and marked the biggest turning point in the landscape of forensic genetics.

IGG evolved as an interdisciplinary practice that combines genomics and computer database technologies, as well as traditional and genetic methods of genealogical research to identify unknown individuals by reconstructing their ancestral lineages and drawing out their family trees. Microarrays that make use of ancestry informative SNPs are the most popular test method available for this analysis. Since its inception, these tests were routinely used as a tool to help search family relationships especially with adoptees and donor-conceived individuals but until 2018, it was never tried with criminal cases to look for the family relatives of unknown perpetrators. However, the principle behind the method is fairly simple; DNA segments exhibiting identity-by-descent (IBD) between the DNA from the crime scene (i.e. belonging to the unknown perpetrator) and an individual in a

genomic database will signal a familial relation between the two. The length of the shared segments can indicate how close the family relationship is. Once distant relatives are identified from the databases, the relevant family trees can be traced to find the closer relatives via genealogical research. When close relatives are within sight, the law enforcement authorities can further narrow down the search using other physical evidence to trace the perpetrator whose DNA signature matches exactly with the crime scene DNA.

IGG first attracted attention of forensic communities after resolving Golden State Killer case in the USA, a 43-year-old case of serial murders and sexual assaults. Since then, IGG had assisted in over 200 old cases, of which at least 28 were solved (Dowdeswell, 2022). The main obstacle currently hindering the use of this technology in criminal identification is the inaccessibility of the large commercial genetic databases for search by law enforcement agencies. Since the technique requires microarray SNP data for the analysis, State owned forensic databases which are primarily based on STR analysis data cannot be used for this purpose. On the other hand, protocols for using IGG methodology in legal setting have not been reviewed yet in full detail by the relevant authorities and the necessary guidelines and the framework for its application have not been established. Further, the details of the IGG work have not yet been scrutinized in court. However, attempts are currently being made to fill these gaps between the experimental setting and the practice and to standardise the method for a more reliable and accurate application within the forensic context.

## RNA: the next paradigm shift in forensic genetics?

The prospect of using RNA data in criminal investigations entails a major new frontier in forensic genetics (Lynch and Fleming, 2021). RNA was previously considered too unusable to be of forensic value. However, recent research has revealed that tiny snippets of RNA can be extremely long-lived. Scientists were able to extract RNA of analysable quality and quantity from 20-year-old dried blood samples, and human remains excavated after four to 44 years. Unlike the DNA complement of an individual that is common for all cells, RNA which represent cells' transcriptome differs depending on the cell type. Consequently, RNA can reveal forensically valuable information that DNA could not, opening a new arena for forensic investigations. For example, RNA profiles can differentiate between menstrual blood and blood found in vessels, a distinction that is important in determining the origin of blood stains. Likewise, different RNA markers have been identified for saliva, semen, and vaginal secretions. In addition, recent research conducted on RNA expression profiles was able to estimate the age of a bloodstain, based on the rate of degradation of different RNA types. Since there is no validated method to determine the age of a bloodstain at present, establishing such a technique will be extremely valuable for forensic scientists. The current RNA analysis technology allows the detection of biomarkers for different body fluids and organs from RNA strands of less than 150 nucleotides and RNA profiling has been carried out for more than 300 forensic cases to complement other DNA based evidence. However, in a forensic point of view, RNA technology is still in its infancy and might need full review of its applicability before its integration into routine casework.

## Expanding boundaries of forensic genetics

The latest developments in molecular techniques, especially the versatility of NGS have expanded the boundaries of forensic genetics to incorporate crimes associated with fauna and flora. For instance, illegal trade of protected or endangered wildlife (wildlife forensics), species mislabelling and adulteration in the food industry and identity forgery or theft of livestock or pet animals can now be investigated using DNA evidence. One of the most critical problems faced during such non-human analysis is identifying the species from which a particular product has originated. These may include cases where a leather watchband or the pulverized remains of animal and plant species present in a medicinal powder need to be tested to verify the presence of parts from protected species. Likewise, there are instances where food and food supplements may carry cheaper or illegal products in contrast to what it claims to contain. On the other hand, the investigation of theft and illegal culling of farm and pet animals may need both taxonomical assertion as well as individualization to trace them to their owner and the perpetrator.

Barcoding and STR analysis have been the methods of choice during the last two decades for forensic genetic investigation at such instances. For example, barcoding analysis using mitochondrial DNA has already been used for meat testing in food safety management and screening for mislabeled, fraudulent or imitation meat in many countries. Likewise, STR panels that were optimized for both species identification, as well as individualization of dogs, cattle and horses, have been made commercially available to ensure wider availability among the forensic community. The establishment of Canine CODIS (Combined DNA Index System) by the American Society for the Prevention of Cruelty to Animals, which archives individual DNA profiles of dogs, helps to identify relationships between

dogs and enables law enforcement agencies to establish connections between breeders and trainers of dog fighting games. Equine-specific STRs analysis is another routinely used protocol for doping control in racing horses. In addition, in-house assays have been developed for the identification of domesticated elephants in countries like Sri Lanka and India. Further, genetic testing of domestic animal biological material has been occasionally used to connect human victims, perpetrators and/or crime scenes during court proceedings providing probative evidence.

More recently, several research studies have emphasized the feasibility of using NGS in the forensic analysis of non-human material based on its inherent advantages over conventional sequencing (Meiklejohn et al., 2021). One such feature is its increased sensitivity which facilitates the recovery of full target sequences even from low quantity and quality DNA samples. For instance, hard matrices such as ivory, teeth, bone and timber contain only a few intact DNA fragments which has made its amplification extremely difficult with conventional PCR methods used in STR or barcoding analysis. However, successful taxonomic assignment of such items carrying highly degraded DNA has been achieved using NGS. Likewise, researchers have applied NGS technology to processed seafood products, to confirm authenticity or detect possible undisclosed allergens such as crustaceans. Another advantage of NGS in wildlife forensics and food authentication involves its capacity to discriminate between the DNA stemming from mixed samples. Since NGS generates individual reads, they can be assigned to separate taxa, whereas a mixed Sanger chromatogram used in conventional barcoding cannot be reliably interpreted. When this feature is coupled with the ultra-sensitivity of NGS, it can be used to detect and characterize taxa present in mixtures at a level low as 1 %. These potentials have led to the development of commercial NGS assays (i.e.,

ThermoFisher Scientific; A38452, A38454) to streamline taxonomic identification of meat and fish in food mixtures for authentication purposes, although such commercial kits are not yet available for wildlife forensics. Further, for taxa in which analysis of multiple regions are needed to ensure reliable and accurate taxonomic identification, new techniques such as "genome skimming" coupled with NGS have demonstrated the ability to streamline the generation of data, for example by reconstructing full mitochondrial genome sequences. It is worth noting, however, that cases from research literature demonstrating what is technically possible in terms of forensic genetics application in non-human sample identification are not adequately validated or compatible with the workflow of forensic casework at present limiting its use in the law enforcement context. Nevertheless, there seems to be great promise for the future in expanding wildlife forensics and food product analysis parallel to the adaptation of NGS technology in forensic labs.

## Concluding remarks

The recent advances in forensic genetics have paved way to resolve both criminal and kinship cases which otherwise would remain mysteries. However, before reaping the full benefit of these modern technologies, several challenges need to be addressed. Many of these challenges are ethical or statutory in nature rather than technical. For example, allowing law enforcement services to access the millions of DNA signatures locked up in private ancestry testing services and making DNA profiling a routine part of the investigation of suspects arrested for major felonies are simple but controversial policy decisions need to be made to facilitate exercising the power of forensic genetics. Initiating a framework for international

collaborations towards establishing forensic DNA databases that comprise STR, microarray and NGS data is another challenge that needs to be met to reap the benefits of the changing paradigms. In addition, the necessity to introduce stringent protocols to fine tune the laboratory guidelines in a standardized framework are necessary to ensure technical and statistical validity of the DNA profile results. This concern is especially valid for those tests conducted based on methods that are not primarily meant to serve forensic investigations such as microarray and NGS. Despite these deficiencies which invite further refinements and research, the changes that took place in the landscape of forensic genetics over the past century hold enormous promise for determining innocence and also to find guilt of offenders. Nonetheless, it is too early to predict whether this change in paradigm in forensic science is powerful enough to significantly discourage criminal offences and make a sustainable upliftment in social security.

## References

Ansorge, W. J. 2009. Next-generation DNA sequencing techniques. *NBiotechnol,* 25**,** 195-203.

Budowle, B., Allard, M. W., Wilson, M. R. & Chakraborty, R. 2003. Forensics and mitochondrial DNA: applications, debates, and foundations. *Annual Review of Genomics Human Genetics,* 4**,** 119-141.

Butler, J. M. 2011. *Advanced Topics in Forensic DNA Typing: Methodology,* London, UK, Elsevier Academic Press.

Chakraborty, R., Stivers, D. N., Su, B., Zhong, Y. & Budowle, B. 1999. The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis,* 20**,** 1682-1696.

de Vries, J. H., Kling, D., Vidaki, A., Arp, P., Kalamara, V., Verbiest, M. M. P. J., et al.Kayser, M. 2022. Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy. *Forensic Science International: Genetics,* 56**,** 102625.

Dowdeswell, T. L. 2022. Forensic genetic genealogy: A profile of cases solved. *Forensic Science International: Genetics,* 58**,** 102679.

Ge, J., Sun, H., Li, H., Liu, C., Yan, J. & Budowle, B. 2014. Future directions of forensic DNA databases. *Croatian Medical Journal,* 55**,** 163-166.

Jeffreys, A. J., Brookfield, J. F. & Semeonoff, R. 1985. Positive identification of an immigration test-case using human DNA fingerprints. *Nature,* 317**,** 818-819.

Lewontin, R. C. & Hartl, D. L. 1991. Population genetics in forensic DNA typing. *Science,* 254**,** 1745-1750.

Lutz, S., Wittig, H., Weisser, H. J., Heizmann, J., Junge, A., Dimo-Simonin, N., et al.Augustin, C. 2000. Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Science International,* 113**,** 97-101.

Lynch, C. & Fleming, R. 2021. RNA-based approaches for body fluid identification in forensic science. *WIREs Forensic Science,* 3**,** e1407.

Meiklejohn, K. A., Burnham-Curtis, M. K., Straughan, D. J., Giles, J. & Moore, M. K. 2021. Current methods, future directions and considerations of DNA-based taxonomic identification in wildlife forensics. *Forensic Science International: Animals and Environments,* 1**,** 100030.

Roewer, L. 2013. DNA fingerprinting in forensics: past, present, future. *Investigative Genetics,* 4**,** 22.

Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1986. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature,* 324**,** 163-166.

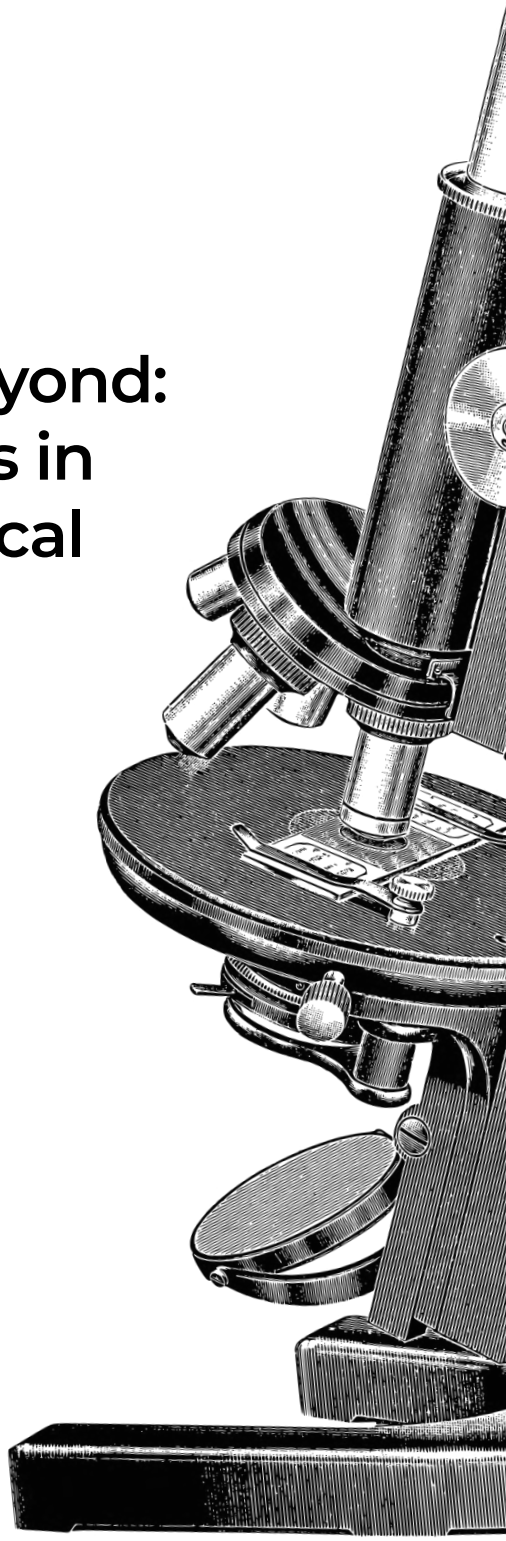
SWGDAM. 2019. Addendum to the SWGDAM interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories to address next generation sequencing. Available: https://www.swgdam.org/_files/ugd/4344b0_91f2b89538844575a9f51867def7be85.pdf [Accessed 15.07.2022].

Yang, Y., Xie, B. & Yan, J. 2014. Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics,* 12**,** 190-197.

Young, J. M. & Linacre, A. 2021. Massively parallel sequencing is unlocking the potential of environmental trace evidence. *Forensic Science International: Genetics,* 50**,** 102393.

Chapter 3

# A journey into a microbial cell and beyond: The power of the lens in shifting microbiological paradigms



W. R. P. Wijesinghe

Department of Botany, Faculty of Science,
University of Peradeniya, Peradeniya, Sri Lanka.
Corresponding author: priyangaw@sci.pdn.ac.lk

## Abstract

Although Thomas Kuhn coined the term ‘paradigm shift’ in 1962, we can find the examples for paradigm shifts in biology throughout the history. When it comes to microbiology, we have to start with the discovery of the microscope itself as this tool expanded the human visualization to the micrometer scale. However, this visualization through a physical lens came about with a number of inherent weaknesses in the system. How the human brain perceived such obstacles in the field of microscopy and the answers were found to get a better image of the microorganisms is discussed here. This pursuit was ended up shifting the paradigm of microscope into a nasoscope challenging some of the earlier perceptions in the scientific thinking. While the physical glass lens and its advancements helped us to see through the microbial cells and cell constituents, a plethora of information of the microbes were hidden in their hereditary information. The early studies to elucidate the nature of genetic material and the heredity depended on the microbiological studies themselves. Although we set the stage of microbiology with the inception of the microscope, as recently as this year, scientists discovered a new group of bacteria that is visible to the naked eye affirming the power of paradigm shifts in biology.

## Keywords

*Genomics, Heredity, Microbes, Microscopy, Nanoscope*

## Introduction

The Greek philosopher Aristotle named his logic as a “syllogism” which proved hugely influential for 2000 years. The core of Aristotle’s system was based on a form of argument from the general to the particular that consisted of three parts, a major premise, a minor premise and a conclusion. Here, as long as the premise is true, then the conclusion is also true. One famous example by Aristotle goes like: all men are mortal, Socrates is a man, and therefore Socrates is mortal. But by the late Middle Ages the enduring authority of the Aristotelian logic was an obstacle to the rise of the physical sciences, which begin with observed facts and then proceed to draw conclusions from those. In this novel paradigm, Francis Bacon and the basis for the scientific method should be commended as a means of inquiring the truth and coming to conclusions based on those observations. Through a process of observations, questioning and experimentation, scientists develop an understanding of cause-and-effect relationships. We call it inductive reasoning and it was a crucial early step towards establishing the scientific method.

Coming to the topic of this chapter however, to observe microbes, one requires a microscope because this new world of organisms is not visible to the naked eye. Hence, the invention of a new tool itself mark the beginning of the new discipline of microbiology. While the human imagination brings the microscopy to the new heights, studying the genetic material of the microbes adds a new dimension to understanding the nature of heredity. It’s fascinating that in the world of paradigm shifts, the discovery of new technologies broadens the horizons of new knowledge and by the same token, the expansion of knowledge leads to the development of new technologies. This virtuous cycle seems to be driving paradigm shifts in an upward trajectory bringing our understanding closer

to the reality of a particular subject. With the help of the following three questions, an attempt has been made to put the paradigm shifts in the context of microbiology looking through a physical as well as a metaphorical ‘lens’.

1. Is microscopy an essential tool to visualize microbes?
2. What was the contribution of microbes to determine the basis of heredity?
3. Are there any bacteria visible to the naked eye?

## Is microscopy an essential tool to visualize microbes?

Here we can equate the meaning of the ‘lens’ to the physical glass lens itself and how it was advanced thanks to the human imagination. However, you would realize that the story we start with a microscope will end up in a nanoscope due to the paradigm shifts took place in the field. Arguably the single most beneficial idea in human health could be the discovery that the microorganisms cause most diseases. The beginning of the germ theory of disease dates back to 1670s when the Dutch lens grinder Anton Van Leeuwenhoek using a single microscope to become the first person to observe microbes. However, while the most literature sources report this story of observing bacteria as the first visual encounter of microbes by the humans, some explicitly argue that it was Robert Hooke but not Leeuwenhoek who discovered the microbes citing the drawings of fungi (*Mucor*) by Robert Hook (Gest, 2004a, 2004b, 2007). Irrespective of whom to give the credit being the first to observe microbes, without doubt they both were geniuses at the time. Evidently this is the beginning of the entire new field of microbiology. Although the microscope allowed microorganisms to be seen as early as the 17$^{th}$ century, it was not until the 19$^{th}$ century the microbiology began to develop as a truly scientific

discipline. This is because it was during the time of the 17th to 19th century, microbiological methods and objectively measuring their activities developed to the point that the scientific method could be applied to study the microorganisms.

## Are microbes living organisms at the first place?

Aristotle was one of the earliest recorded scholars to put forward the theory of spontaneous generation i.e., life can arise from non-living matter (Zwier, 2018). After the Francisco Redi's experiment on meat, it was disproved the fact that macro-organisms do not arise spontaneously. However, the theory of spontaneous generation of microorganisms remained a viable idea throughout the 18th century and into the early 19th century making the point that once an idea is normalized according to Kuhn, the believers of it are hanging by the fingernails not to abandon the 'normalcy' to adopt the new findings. After several decades, it was Louis Pasteur who discredited the theory of spontaneous generation of microbes with the help of his swan-necked flask. With this contribution by Pasteur and his follow up demonstrations on microbes marked the birth of microbiology as a science. Moreover, by comparing the samples taken from vats with good wine and sour wine, Pasteur observed budding yeast cells in the vats of good wine and rod-shaped bacterial cells in the vats of sour wine. Now we will consider the technology that developed to observe such yeast and bacterial cells.

## Light microscopy

Light microscopy uses visible or ultraviolet light to illuminate an object. The light passes through several glass lenses that alter the path of the light

and produces a magnified image of the object. Here, visible light is bent by a series of glass lenses to achieve magnification, and this physical phenomenon is known as refraction. However, magnifying an image using a microscope is only useful if details of the object can be accurately preserved in the image. The resolution is the degree to which the details in the specimen is retained in the magnified image and therefore of paramount importance. Although lenses marked the beginning of observing microbes, lenses used in microscopes have several inherent problems that distort the magnified image. The following brief note indicates how scientists overcame such issues with the ingenuity of the human thinking.

Contrast is a property that requires to visualize an object from the surrounding background. Microorganisms are mainly composed of water and hence these poses difficulties viewing microbes due to the inability to distinguish cells from background. Staining the microbes with a colored compound increase the contrast. Scientists figured different ways of staining microbes from simple staining with a single dye to differential staining procedures to stain specific type of microbes or particular structures of a microorganism. For example, one could visualize the live and dead bacterial cells in the same culture by a combination of SYTO9 to stain the live cells while propidium iodide to stain the dead cells.

An array of microscopes has been developed for varying applications in microbiology. In a bright-field microscope, visible light is transmitted through the specimen. The specimen generally appeared dark on a bright background. This microscope usually requires staining of specimens and rarely used to observe live microbes. In contrast the dark-field microscope has been designed to eliminate the need for staining the specimen. The dark-field condenser lens focuses light on the specimen at an oblique angle, such that light does not reflect off an object and does not enter the objective

lens. Therefore, the field will appear dark while microorganisms will appear very bright on a dark background.

The fluorescence microscopes are designed in such a way that the specimen can be illuminated at one wave length of light and observed by a light emitted at a different wavelength. This allows the use of multitude of fluorescence dyes to stain the microbes. The phase-contrast microscope has designed to view microbial structures without staining. Here, light that passes through a cell or a cell structure is slowed down relative to the light that passes directly through the less dense surrounding medium. The confocal scanning microscope is a different type of microscope that does not form a two-dimensional optical image of the specimen as occurs in a conventional microscope. The design of the confocal scanning microscope totally eliminates the diffracted light that tends to blur the image of a conventional microscope.

## Electron microscopy

While light microscope uses visible or ultraviolet light, the electron microscope utilizes a beam of electrons that permits much greater resolution and thus much higher magnification than the light microscope. The greater resolution of the electron microscope is possible because the wavelength of an electron beam is much shorter than that of light in the visible range of the electromagnetic spectrum. Consequently, the magnification is in excess of 100,000×. Therefore, the electron microscopy provides sufficient magnification and resolution to observe viruses and internal structures of microorganisms. There are two major types of electron microscopes and the Transmission Electron Microscope (TEM) uses an electron beam that passes through the specimen while Scanning

Electron Microscope (SEM) uses an electron beam that is scanned across the surface of a specimen. However, the electron microscopy requires sample preparatory measures that eventually kill the microbe. In this context, is there a way to visualize microbes at nanoscale without killing them. Apparently, scientists have shifted that paradigm as well in what's called Super-resolution fluorescence microscopy and we will focus on that in the next section.

## Super-resolution fluorescence microscopy – where optical microscope becomes a nanoscope

In 1873 Ernst Abbe published an equation (after all it is an equation) demonstrating how resolution of the microscopy is limited by the wavelength of the light (Abbe, 1873). The 20$^{th}$ century scientists grew believing that they would never be able to observe objects smaller than roughly the half the wavelength of light. By the same token we used to memorize this smallest value as 0.2 micrometers as the right answer for our exams. Abbe's equation may still hold true but, is there a way to bypass it. Three scientists Eric Betzig, Stefan W. Hell and William E. Moerner broke the above barrier with the help of fluorescent molecules and single molecule microscopy and won the Nobel Prize in Chemistry in 2014. How the Abbe's diffraction limit was bypassed can be rewarded to two major principles.

In 1994, Stefan Hell published an article proposing his idea called stimulated emission depletion (STED) (Kner et al., 2020). Here, a light pulse excites all the fluorescent molecules, while another light pulse quenches fluorescence from all molecules leaving aside a nanometer-sized volume. Only this volume is then registered. Now, by sweeping along the

object and continuously measuring light levels, it's possible to get the full image. In year 2000 using his STED microscope, Stefan Hell demonstrated his theoretical idea practically by imaging an *E. coli* cell at a resolution that was never possible with an optical microscope. In contrast to the STED, the second principle lies on the single-molecule microscopy that developed based on independent contributions by Eric Betzig and W.E. Moerner.

W.E. Moerner was the first to detect a single fluorescent molecule (Moerner & Kador, 1989). In most of the analytical methods, scientists analyze the average of millions of molecules via absorption or fluorescence measurements. Researchers had to accept this as the norm because nothing else was possible. However, people dreamt of measuring single molecules. In 1989, W.E. Moerner for the first time was able to measure the light absorption of a single molecule. This pioneering work opened up the door to a brighter future for single molecule experiments and a newer future for many chemists to pay attention in this new direction. Eric Betzig was one of them and we will consider his contribution next.

Just like Stefan Hell, Eric Betzig was pondering the ways in which to bypass the Abbe's diffraction limit. Even after quitting his job at the Bell Labs, one cold Winter Day, a new idea came to his mind; would it be possible to overcome the diffraction limit by using molecules with different properties i.e., molecules that fluoresce with different colors. In fact, Eric Betzig had already detected fluorescence in single molecules using near-field microscopy. He now began to think of a way a regular microscope could yield the same resolution if different molecules glowed with different colors. The idea was to register one image pertaining to a single color at a time. If all molecules of one color were dispersed and never closer to each other than the Abbe's diffraction limitation (0.2 micrometers), their positions could be determined precisely. Once such images for different

colors were superimposed, the complete image would get a resolution even if their distance was just a few nanometers apart. This lays the foundation to circumvent Abbe's diffraction limit.

The super-resolution fluorescence microscopy that came to fruition as a result of circumventing the Abbe's diffraction limits was truly a paradigm shift against ideology by even a famous scientist like Erwin Schrödinger, who once proclaimed "we never experiment with one electron, atom or small molecule. How can we use single molecule labels to surpass Abbe's optical diffraction limit, a fundamental physical effect in the far-field?".

## What was the contribution of microbes to determine the basis of heredity?

Here we have to broaden the meaning of the world 'lens' to incorporate the metaphorical meaning of the word that suggests a vision for understanding something new. This vision into the hereditary material with the help of microbes has contributed to bring the field to its pinnacle. Long before scientists understood its mechanisms, human beings had some understanding on genetics, the science of heredity. For example, as early as 5000 B.C. civilizations around the world crossbred livestock and crops like wheat and corn to produce superior new varieties. However, in the process of delineating the real mechanics of heredity, it seems human imagination was deceptive from time to time. A classic example was around 300 years back, the visualization of a tiny man inside a sperm by Nicolaas Hartsoeker in line with the concept of preformationism.

However, by the 19$^{th}$ century a scientific monk, Gregor Mendel got the human imagination with respect to the heredity on the right track. Mendel observed the ways in which traits like height, color, and wrinkled surfaces

of seeds in the pea plants passed from one generation to the other. He observed that most such traits are carried by factors or now what we call genes. He also determined that the genes could be dominant or recessive and also some fundamental rules of their patterns of inheritance. What Mendel described in abstract manner was truly remarkable at the time but it took another century to understand the true nature of the genes. In 1941, George Beadle and Edward Tatum studying on the fungus *Neurospora* established that specific segments of DNA, called genes encoded the information for making specific proteins. Their studies supported the one-gene, one-enzyme hypothesis although now we know that it's not true at all the times. They also introduced changes (mutations) by exposing *Neurospora* spores to x-rays. By 1958, George Beadle, Edward Tatum together with Joshua Lederberg shared the Nobel prize on microbial genetics, a fascinating discipline that contributes immensely to science even to date and years to come.

## DNA is the hereditary molecule

Even in the middle of the 20$^{th}$ century, some giants of the scientific community believed that DNA was a 'boring' or 'stupid' molecule whereas proteins should be the molecule endowed with heredity. Nevertheless, a group of scientists visualized an avenue to make their case that it's the DNA that acts like the hereditary molecules. The property that they utilized was the abilty to transform some non-disease-causing strains of the bacterium *Streptococcus pneumoniae* into a virulent strain that causes pneumonia. In the late 1920s Fred Griffith observed that disease causing strains of *S. pneumoniae* produces a polysaccharide capsule whereas avirulent strains do not. Mice infected with the capsulated strain died even with minor dose

while non-capsulated strains did not cause death. When Griffith injected mixtures of heat-killed capsulated and live non-capsulated strains into mice, the mice died and it was also possible to isolate capsulated strains from such mice.

The experiment done by Avery, McCarty and MacLeod provided the molecular explanation for this event in 1941 by separating the different types of molecules in the debris of the dead encapsulated *S. pneumoniae* cells and testing each molecular class for ability to cause transformation. They delineated the fact that it's the DNA that contains the genetic information based on the following meticulous steps. At the very beginning they removed all the remaining fragments of the bacterial coat from the debris. They dissolved the lipids in alcohol but there was no difference in transformation. They then stripped off proteins with chloroform and still the transforming activity was retained intact. Next, they digested different proteins with different enzymes and still the transformation ability prevailed. Heating the mixture to 65 ˚C to denature most proteins did not alter the transformation either. The only way to kill the transformation was to digest the material with an enzyme that degrades DNA. They also tested the transforming principle with UV light, electrophoresis and chemical analysis. The transforming material should be DNA and it can induce predictable and hereditary changes in cells Avery concluded.

Once more and more evidence was accumulating from other labs too in line with DNA being the transforming molecule but not proteins, even the strong skeptics had to convert to the believers of DNA. This exemplifies the 'crisis stage' in the Kuhn's cycle of paradigm shifts where voluminous data to support DNA as the molecule of heredity accumulates amidst inability to explaining the transformation principle with any other biomolecule including proteins. Here we can also understand the

importance of the study of microbes in elucidating the nature of nucleic material. Without the transformation principle and the associated microbial work how long would it take for the scientists to determine the biomolecule of heredity is a question worth evaluating. For example, can we predict the understanding of the nature of nucleic acids without the knowledge of the transforming principle? If it was the case, what path would have been taken to come to the present understanding of the nucleic acids by the scientists?

## DNA sequencing

Two times Nobel Prize winner Fred Sanger in mid 1970s understood the importance of DNA sequencing i.e., the determination of the order of A, G, C, and Ts in the genome of an organism. The method developed by Sanger was capable of handling a small virus genome like ϕX-174 having only a 5386 base pairs. However, applying the same sequencing method even after a number of technological advancements, it took 3 billion US dollars and 13 years to sequence the first human genome. Now it is possible to sequence the entire human genome under $ 1000 within few hours. Moreover, current projects are embarking on sequencing millions of human genomes in the world thanks to the advancements in the field. Similar projects convince us that there are more microbial cells in the human body than the human cells themselves. There was a quantum leap when scientists understood the capabilities of second-generation or the next-generation sequencing compared to the Sanger sequencing.

## Second-Generation Sequencing/ Next-Generation Sequencing (NGS)

Sanger sequencing is based on a single source of DNA. This means it can determine the sequence of A, G, C and T coming from a single entity of pure DNA in a sample. For example, if you are using a plasmid DNA to sequence, your DNA sample should contain only the plasmid of your interest. If the plasmid DNA sample is contaminated with another plasmid or genomic DNA of the bacteria, the Sanger method will fail and it will give lots of 'N's instead of normal four nucleotides as the readout. Likewise, if the DNA of interest is a PCR product, Sanger method will determine only one amplicon at a time. That amplicon can be an amplified product of a single bacterial gene. If you want to sequence different DNA (e.g., more than one gene), they have to be run separately.

In contrast to Sanger sequencing, Next-Generation Sequencing (NGS) can use the DNA from the entire bacterium comprising the genome as well as the plasmid as the source of DNA to begin the sequencing process. Because of this capability, another term used for NGS is Massively Parallel Sequencing (MPS) meaning it can process millions of sequencing reactions in parallel. Although MPS is more relevant in explaining the mechanics of the technology, the term was not popular as NGS. With this new potential of NGS, now the sequencing can be done in high-throughput fashion applying to genome-wide as well as population scale. Taking together, NGS can be considered as a paradigm shift that revolutionized nearly every biological discipline including microbiology with its universal applications.

Although not in the limelight, Nick McCooke could be considered as the mastermind behind the NGS. His team sequenced the complete genome of bacteriophage ϕX-174 using a method called sequencing by synthesis (SBS),

the same genome Fred Sanger first sequenced using so called dideoxy termination method. However, the SBS technology generated significantly more sequence data, delivering over 3 million bases from a single run.

DNA sequencing and related omics studies are continued to marvel the scientific as well as the lay community on daily basis. The study of microbial community DNA (of a soil sample, ocean water or entire gut bacteria at once) as opposed to pure cultures, gave birth to a new discipline called metagenomics. In medicine, a breakthrough liquid biopsy like Karius Test® is now capable of non-invasively and rapidly detecting over 1000 pathogens from a single blood sample causing both deep-seated and bloodstream infections. Here, the analysis of the entire pathogen is not a must but the DNA shed from the microbe and circulate in the blood is sufficient to determine the identity of the pathogen based on its species-specific nature or the unique sequence of the DNA fragments.

## Are there any bacteria visible to the naked eye?

Bacteria are by definition microscopic. However, as recently as June, 2022 scientists reported a giant bacterium discovered from Caribbean mangroves that grows up to a size of a peanut (Pennisi, 2022; Volland et al., 2022). Biologist Olivier Gros first observed this organism in 2009 while exploring the mangroves of Guadeloupe, where he works at the University of the Antilles in the French West Indies. However, in 2018 marine biologist Jean-Marie Volland at Lawrence Berkeley National Laboratory in California took a fresh look at the bacteria using a range of methods, including transmission electron microscopy and fluorescence *in situ* hybridization. In this way, he was able to confirm that it was a single cell bacterium but not a fungus although Olivier Gros once used to think about the giant species back in the 2009.

This sulfur oxidizing, carbon fixing bacterium is 5000 times bigger than the most bacteria. Its threadlike single and sessile cell could grow up to two centimeters and hence the authors proposed the name *Thiomargarita magnifica* to name the newcomer. The authors analyzed the membranes of *T. magnifica* using osmium tetroxide or the fluorescent dye FM 1-43X, and visualized entire filaments in three dimensions with hard x-ray tomography and confocal laser scanning microscopy. Filament sections that ran up to 850 mm in length were visualized with transmission electron microscopy (TEM). All these techniques consistently showed that each filament was one continuous cell for nearly its entire length without constrictions or division septa.

More importantly, the nuclear material of the bacterium is enclosed inside a sac. Ribosomes are also present in this membrane-bound organelles. These membrane-bound compartments that contain DNA clusters were coined by the authors as “pepins” due to their resemblance to small seeds in fruits. There are hundreds of thousands of such pepins present in a cell. DNA is also extraordinary in this bacterium as there are 11 million base pairs comprising over 11,000 clearly identifiable genes. Furthermore, very high polyploidy of about 500,000 copies of the genomic material was observed inside a cell. Another remarkability of this finding is that *T. magnifica* has a genome as large as the baker’s yeast *Saccharomyces cerevisiae* (12.1Mb) and contains more genes than the model fungus *Aspergillus nidulans* (~ 9500 genes).

The above extraordinary size of the bacterium challenges certain bioenergetic and biophysical limitations on cell growth. Similarly, the compartmentalization of genomic material and ribosomes in translationally active organelles bound by bioenergetic membranes, dimorphic developmental cycle in which genome copies are asymmetrically

segregated into daughter cells indicating a gain of complexity in the *Thiomargarita* lineage are all contribute to challenge the traditional concepts of bacterial cells.

Authors speculate that changes in spatial organization of cellular components, such as DNA and ribosome compartmentalization and rearrangement of the bioenergetic membrane system, may allow *T. magnifica* to overcome many such limitations. The research team used single cell genomics to analyze five of the bacterial cells collected from a single sunken leaf in the mangrove. In parallel, using a labeling technique known as BONCAT to identify areas involved in protein-making activities, it was confirmed that the entire bacterial cell was active.

## One giant bacterium, multiple fundamental questions

The authors raise multiple new research questions based on their finding. Among them, one is what is the role of the bacterium in the mangrove ecosystem? Another question is whether the new organelles named pepins play a role in the evolution of the *Thiomargarita magnifica* extreme size, and whether or not pepins are present in other bacterial species. The precise formation of pepins and how molecular processes within and outside of these structures occur and are regulated.

If we see this example through the lens of evolution, the biology keeps us surprising bringing exceptions to the existing norm. On the other hand, phenomena like ‘plasticity’ and ‘epigenetics’ or even the ‘eye-sight’ seem to be already challenging the classical explanations of evolutionary principles like natural selection. Therefore, whether it’s a discovery of a new species like *T. magnifica* or even challenging a bedrock principle like the theory of evolution with a concept like ‘epigenetics’, we have to take a

step back sometimes and rethink that all these principles were out there all along all the time, but it's the human perception that takes twists and turns to fathom the paradigm shifts in biology or elsewhere.

Following the theme of this book chapter, here with respect to the giant bacterium, we don't have much sense to the word 'lens' according to the context we began with i.e., we set the stage for the birth of a new discipline called microbiology that totally depended on a microscope. Now that scientists discover microbes of which visualization does not depend on a microscope and such is the case when we are dealing with paradigm shifts.

## Concluding remarks

Weighing the above facts and figures of the chapter, we can make some inferences regarding the paradigm shifts in biology. If a technology is involved as in the case of microscopy or DNA sequencing above, new advancements of the technology are taking place to broaden the horizons of the human senses. For example, micrometer resolution became nanometer while the initial observations of the nuclear material at the level of chromosomes narrowed down to the level of single nucleotide resolution.

The frameworks we set to define a particular discipline or area has to be restructured from time to time in the presence of novel findings. For example, demarcation of the discipline microbiology took place only after the discovery of microscopes. In that context bacteria fell in the middle of the microbial spectrum among other microbes. However, now we are learning that the existence of bacteria which are visible to the naked eye. The following quote is coming from the United States former secretary of defense Donald Rumsfeld, that goes "There are known knowns - these are

things we know that we know, there are known unknowns - that is to say, there are things that we know we don't know. But there are also unknown unknowns - there are things we don't know we don't know". The above statement coming from the political world implies that if anything lies in the third category it leaves us clueless, however, the paradigm shifts are the segues that give us glimpses into unknown unknowns in the scientific arena.

## References

Abbe, E. 1873. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv Für Mikroskopische Anatomie*. https://doi.org/10.1007/bf02956173

Gest, H. 2004a. The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, fellows of the Royal Society. In *Notes and Records of the Royal Society*. https://doi.org/10.1098/rsnr.2004.0055

Gest, H. 2004b. The discovery of microorganisms revisited. In *ASM News*.

Gest, H. 2007. Fresh views of 17th-century discoveries by Hooke and van Leeuwenhoek. In *Microbe*. https://doi.org/10.1128/microbe.2.483.1

Kner, P., Manley, S., Shechtman, Y., & Stallinga, S. 2020. 25 th Anniversary of STED Microscopy and the 20 th Anniversary of SIM: feature introduction . *Biomedical Optics Express*. https://doi.org/10.1364/boe.391490

Moerner, W. E., & Kador, L. 1989. Optical detection and spectroscopy of single molecules in a solid. *Physical Review Letters*. https://doi.org/10.1103/PhysRevLett.62.2535

Pennisi, E. 2022. Mammoth mangrove bacterium has complex cell. In *Science*. https://doi.org/10.1126/science.adb1758

Volland, J.-M., Gonzalez-Rizzo, S., Gros, O., Tyml, T., Ivanova, N., Schulz, F., Goudeau, D., et. al. 2022. A centimeter-long bacterium with DNA compartmentalized in membrane-bound organelles. *BioRxiv*.

Zwier, K. R. 2018. Methodology in Aristotle's Theory of Spontaneous Generation. *Journal of the History of Biology*. https://doi.org/10.1007/s10739-017-9494-7

# Chapter 4

# Quantum biology: A novel direction in biological sciences

D. V. S. Kaluthanthri[1], P. N. Dasanayaka[1]
S. A. C. N. Perera[2]*

[1]Department of Botany, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka
[2]Department of Biology, Faculty of Agriculture, University of Peradeniya, Peradeniya, Sri Lanka
* Corresponding author: chandrikaperera@agri.pdn.ac.lk

## Abstract

Quantum mechanics includes the fundamental theory that describes the properties of subatomic particles. Applications of quantum mechanics in numerous biological phenomena including the origin of life, DNA mutations, photosynthesis, enzyme catalysis, magnetoreception, olfaction and cognition have been described so far, even to the point of presenting controversial ideas in some cases. Apparently, the novel knowledge on the non-trivial role of quantum mechanics in biological processes will help in the development of novel technological advancements such as quantum computers, artificial photosynthetic systems, olfactory sensing devices and artificial neural networks. The foundation for such advancements has already been established. Ultimately, the advances in quantum biology may lead to an upgrade of current technology and maybe even to the establishment of life on other planets. In this chapter, the current status of quantum biology, its applications and future directions will be discussed.

## Keywords

*Coherence, Entanglement, Quantum biology, Radical pair, Tunnelling*

## What is quantum biology?

Biology is a science which describes the functioning of natural systems from replication of genetically coded information to operation methods of multistep, complex reactions. Quantum mechanics is one of the major domains of physics which includes the uncertainty principle, wave–particle duality, quantization of energies and the modification of classical probability laws. Over the past few decades, scientists have been studying the link between quantum physics and biology (Abbot et al., 2008; Ball, 2011; Collini et al., 2010; Gerlich et al., 2011; Lambert et al., 2013; Lloyd, 2011; Panichayangangkoon et al., 2010; Romero-Isart et al., 2010; Vedral, 2011). Understanding the ways of surviving the quantum effect by living organisms in environments which are 'wet, noisy and warm' etc. has been a major focus in such research (Fleming et al., 2011). Even though there is nothing new in the occurrence of quantum effects in biological systems, the revealing of non-trivial size and temperature range that the quantum effects can occur in a given biological system have contributed to a paradigm shift in biology. Accordingly, quantum biology has become a novel field of science which can account for understanding biological processes where classical laws of physics cannot be precisely applied.

## Importance of quantum biology

Biological sensors possess numerous features that cannot be found in fabricated devices so far. Out of them, high selectivity and sensitivity, robustness in noisy environments, compactness and room temperature fabrication from readily available materials are highlighted. Biochemists and molecular biologists have studied and described almost all biological functions in terms of the arrangements and rearrangements of classical

molecular units prior to the era of quantum biology. In most cases, the attempts to emulate biological sensors through artificial generation of the structure and chemistry have failed. Engineers who tend to develop devices mimicking biological sensors require knowledge to exploit quantum effects and for this reason, the field of quantum biology has been proven to be highly useful, and has gained demand.

There are two fundamental considerations in understanding the quantum level influence in any biological system. These questions namely are; how does the quantum effect sustain in a complex, hot and noisy biological environment? and, how does the quantum effect play its non-trivial role in such important biological processes? The higher the complexity of any given biological system, tackling these two questions would become more difficult. In order to reveal and understand the unknown factors and the essential ingredients in biological systems affecting the quantum dynamics, there should be a clean platform and quantum simulation is one of such platforms (Georgescu et al., 2014).

## Applications of quantum biology

The influence of quantum mechanics in biological systems can be divided into three classes. Of them, class one includes trivial influences such as quantum mechanics dictating energies and molecular orbitals. The class two is molecular dynamics and chemical kinetics including ultra-fast molecular transitions through conical intersection and chemical reactions involving electron and proton tunnelling. There are several possible examples such as olfaction, magneto-reception in birds and photosynthetic light harvesting which can be included in the last class; i.e. functional necessity. In this chapter, seven candidates for quantum biology including**;**

the origin of life, DNA mutations, photosynthesis, enzyme catalysis, magnetoreception, olfaction and cognition will be discussed. Out of them, photosynthesis and enzyme catalysis can be categorized as transport processes whereas magnetoreception, olfaction and cognition are grouped under sensing processes. The potential quantum processes in several biological phenomena which will be discussed in the chapter are summarized in Table 1.

**Table 1:** potential quantum mechanical processes in biological phenomena.

| **Biological phenomenon** | **Quantum process** |
|---|---|
| Origin of life | Quantum superposition |
| | Coherence |
| DNA mutations | Coherence |
| | Proton tunnelling |
| Photosynthesis | Quantum tunnelling |
| | Coherence |
| Enzyme catalysis | Quantum tunnelling |
| Magnetoreception | Radical pair entanglement |
| Olfaction | Quantum tunnelling |
| Cognition | Coherence |

## Origin of life

Every single living cell can be considered as a quantum processor that is comprised of specific genetic information having almost four billion years of evolutionary history. There are debates regarding the way these unique information contents and information processing mechanisms existed in the first place (Davies, 2003). Even though quantum mechanics can explain the

atomic structure of the genetic code, and all other molecules that can be found in living organisms, there is no sufficient basis for explaining the origin of the genetic code and specific translated proteins. Nevertheless, the question of how the ancient soup of classical molecular building blocks randomly discovered the appropriate combination in a reasonable period of time to convert the non-living into living state has gained much attention within the scientific community. Apparently, given a pre-biotic soup composed of all the matter in the universe, it should have taken a much longer time than the age of the universe to form a single protein by chance. This is why the role of quantum mechanics in the origin of life should be non-trivial. Since a quantum system can exist in superpositions of states, searching for the appropriate combination necessary for the conversion of non-living into living state may proceed much faster (Farhi and Gutmann, 1998). However, there should be a teleological aspect of all searches with prior knowledge regarding the target which is the living state. But the ancient soup of classical molecular building blocks could not have such knowledge. Instead of searching for the living state, searching for the replicator which has the ability for making copies itself may sound quantum mechanically more meaningful. McFadden (2001) stated that replicating molecules impact their environment much more strongly than a random molecule in a quantum soup. After the existence of life, living systems may have had the property of being discovered in the subsequent quantum searches for the living state.

## DNA mutations

Biological molecules serve as either specialized chemicals or informational molecules reflecting the underlying dualism between phenotype and

genotype. The informational molecules are DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) of which the structure can be explained using quantum mechanics. According to Darwinian theory, cells individually undergo mutations, and the process of natural selection determines whether these mutations are beneficial to the organism or not, rather than DNA molecules actively participating in this selection process. Contrasting to Darwinian theory, quantum hypothesis describes the DNA molecules as quantum computers which actively engage in the determination of which mutation is the most beneficial.

As a result of the development of the subjects of quantum computation and quantum information processing, the classical information "bit" was replaced by its quantum counterpart, the "qubit." However, as it is assumed that the informational molecules store and process classical bits rather than qubits, the involvement of quantum mechanics in an information-processing role, is disregarded in most cases. However, there are a number of facts supporting the claim that this assumption may be wrong in some circumstances. For instance, quantum mechanics and quantum information processing have played a significant role in the emergence of life from non-living chemical systems and quantum information processing have had a sporadic role in further development process of life. Apparently, life has evolved some form of quantum behaviour as a refinement.

Even though quantum information processing is more efficient than classical information processing, its application is greatly limited by the requirement of quantum coherence. Quantum coherence refers to the capability of a quantum state to maintain its superposition and entanglement under the effect of interactions and thermalization (Konik, 2021). In other words, coherence is the phenomenon of multitasking of quantum entities where an object behaves like a wave so that it can follow

multiple pathways simultaneously. In a practical aspect, interactions between the quantum system and its environment serve to decohere the wave function and this is the major challenge in developing quantum computers.

However, DNA molecules can remain in coherent state for a short time (femto seconds to picoseconds only). McFadden (2001) claimed that mutations might occur as a result of quantum fluctuations. The structure of nucleotide bases can be spontaneously altered by proton tunnelling resulting in incorrect pair bonding. In some circumstances, the genetic code is considered as a quantum code where the superpositions of coding states might occur resulting in spontaneous errors in base pairing (McFadden and Al-Khalili, 1999).

## Transport processes: Photosynthesis, Enzyme catalysis

### *Photosynthesis*

Photosynthesis is a biological process which generates fuel for various activities through conversion of light energy into chemical energy. Inside a photosynthetic cell, this process is initiated with the absorption of photons by light harvesting complexes such as chlorophyll, and proceeds through transferring of energy in the form of electronic excitation to the reaction center where charge separation is driven (Fleming et al., 2011). The overall photosynthetic process is not thermodynamically efficient, as it has less than 10% conversion capacity of light to biomass (Zhu et al., 2008). However, on the basis of quantum mechanics, the photon capturing and transportation process is highly efficient and remarkably robust (Chen et al., 2013). The development of next-generation green energy solutions and optoelectronics may depend on accurate revealing of operative

mechanisms. The vital role of quantum dynamics in photosynthetic reaction has been revealed through both theoretical and experimental studies. The capturing of photons, which is the quantum of light, by light harvesting molecule such as chlorophyll knocks out an electron from light harvesting molecule and produces a quantum particle called an exciton. This exciton travels to the reaction centre through the surrounding molecules. For the maximum efficiency, the excitons should reach the reaction centre with minimum energy loss. In order to minimize the energy loss, excitons should reach the reaction centre *via* the most efficient route through the surrounding molecules. For that, the excitons should reach the reaction centre at high-speed, finding the most efficient path out of all possible paths. This is where quantum coherence is used by the exciton to travel through all the available paths simultaneously through the super position of states, to reach the reaction centre (Collin et al., 2010). This process is referred to as 'quantum walk'.

The occurrence of strong coupling between excitons and the consequent delocalization of their excitation is a quantum effect which is capable of modifying energy transfer pathways due to modifications of their collective dipole configurations. Even though it is more conjectural, quantum entanglement has also been proposed to play a role in this system (Sarovar et al., 2010).

Vibrational coherence lasting in the picosecond scale has been detected in bacterial and plant light harvesting complexes using femto second transient absorption spectroscopy method (Agarval et al., 2000; Chachisvilis et al., 1994; Kumble et al., 1996; Novoderezhkin et al., 2000; Novoderezhkin et al., 2004; Vos et al., 1994). Moreover, a two-dimensional electronic spectroscopy (2D-ES) method has been used to monitor the decay of coherent superposition in light harvesting complexes revealing the

presence of cross peaks that oscillated in time (Brixner et al., 2005). The presence of quantum coherence with respect to light harvesting complexes of green sulfur bacteria (Fenna-Matthews-Olson/ FMO complex) and higher plants (LHCII complex) have been reported by Engel et al. (2007) and Schlau-Cohen et al. (2009), respectively.

### *Enzyme catalysis*

The enzymes are proteins that catalyze biochemical reactions. They enhance the rates of biochemical reactions through lowering the free energy activation barriers. With enzymes, the rates of biochemical reactions can be accelerated by large factors in a way which cannot be accomplished through any other conventional catalytic events. Without the presence of a specific enzyme, the same reaction would occur at the slowest speed which would take a very long time estimated to be about half of the earth's age in years (Frick et al., 1987). Furthermore, the enzymes provide directions to the reactions through conformational changes to stabilise the transition state of the substrate by electrostatic effect. In classical mechanics, the substrates with higher thermal energies than the activation energy, would pass over the barrier, whereas substrates with lower thermal energies would fail to do so. The mechanism of reaction rates which are temperature independent and proceed under low temperature cannot be well explained using the principles of classical mechanics. This is where the concept of quantum tunnelling plays an important role in most enzyme driven reactions (Doll and Finke, 2003). Quantum mechanical particles such as electrons and protons passing through rather than over a potential barrier with a height greater than the total energy of the particles, is called quantum tunnelling. Here, the quantum particles do not have to overcome the energy barrier. They can simply tunnel through the barrier. When

quantum particles reach one side of the barrier they instantaneously can reappear on the other side of the barrier as they can exhibit wave-like behaviour. Currently, most scientists believe quantum tunnelling to play an essential role in enzymatic reactions. However, there are contradictory views on whether quantum tunnelling is involved directly in enzyme catalysis.

Quantum tunnelling has been reported to take place in the metamorphosis of amphibians when the tadpole breaks down and reassembling a frog (Gross and Lapiere, 1962). Interestingly, in the metamorphosis of amphibians, the tail is broken down into cells and redistributed as proteins including enzymes and collagens in the newly forming frog. Here, the strong bonds between collagen proteins of the tadpole's tail are broken down with the aid of enzymes at an accelerated rate. This incredible acceleration may be due to the quantum tunnelling of responsible enzymes. However, it is still controversial whether this process is specifically affected by quantum tunnelling or not (Delgado, 2017).

Another debatable question is whether the enzymes have evolved with the quantum tunnelling capacity in natural selection. Interestingly, the basis for the selection throughout the evolutionary process of electron transfer proteins, which are generally called as oxidoreductases, is the proximity of the redox centres in electron transfer chains (Moser et al., 2006). More evidence supporting the evolution of quantum tunnelling capacity of enzymes are currently emerging (Klinman and Kohen, 2014).

## Sensing: Magnetoreception, Olfaction, Cognition

### *Magnetoreception*

A number of animal species including fish, insects, amphibians, mammals, reptiles and birds have the ability to obtain directional and positional

information from the Earth's magnetic field (Wiltschko and Wiltschko, 2005). There are two types of avian magnetic compasses. The first one can be described using a classical mechanism. Here, the magnetic receptors that can be found in the upper beak of birds such as homing pigeons, screen the intensity of a magnetic field and gather information on position rather than direction (Wiltschko and Wiltschko, 2006). The second type of compass can be found in some birds such as the European robin in which the cryptochrome located in the bird's eye involves in the magnetoreception. In this case, only the magnetic field lines that are inclined upwards or downwards are pointed out. Such an avian compass cannot discriminate between magnetic South and North and is only activated in blue or green light. The mechanism behind such compasses can be explained using quantum mechanics. Apparently, there are two types of molecules as donor and acceptor molecules which are involved in this process. As the donor molecule gets excited by light absorption, an electron is transferred to an acceptor molecule resulting in both molecules having an unpaired electron which begins to spin. Such a pair is called a radical pair. There are two possible states that can be exhibited by a radical pair, namely a singlet or a triplet state, which are entangled. In the singlet state, spins of the electrons are anti-parallel whereas in the triplet state, electron spins are parallel. The radical pair undergoes coherent quantum oscillations between these two states. The rate of this process depends on the orientation of the host molecule with respect to the Earth's magnetic field (Kominis, 2012). According to Ritz *et al*. (2010), there are weak magnetic fields enhancing cryptochrome responses in some plants.

### *Olfaction*

Olfaction system facilitates the sense of thousands of molecules in living organisms. At first, this system was assumed to function by following a lock and key model. According to this explanation, different types of odorant molecules bind to different types of olfactory sensors and the sense of smell entirely depends on the level of affinity of odorant molecules to different odorant molecules. The actual reason behind this difference in sensing level cannot be explained by the lock and key model. Furthermore, the lock and key model cannot be used to explain the differently sensing mechanisms of the different odorant molecules with similar shape and affinity levels. This is why the explanation of quantum mechanics based on electron tunnelling gained higher validity in explaining the theory behind olfaction (Turin, 1996). The nose or the body part which is involved in smelling can be identified as a vibrational spectrometer. When an odorant has the matching vibrational frequency with the olfactory receptor (sensor), binding of odorant and receptor occurs. Thereby, the receptor gets activated and causes electron tunnelling between different energy states in the receptor. As a result, an electrical signal corresponding to the odorant molecule's vibrational signal is generated and transduced to the brain for recognition (Brookes et al., 2007).

### *Cognition*

Currently, the underlying mechanisms behind the brain functions are widely discussed. There are a number of speculations regarding the quantum mechanical role in the ethereal process of consciousness (Lockwood, 1989; Stapp, 2009). Out of them, the hypothesis of quantum-mechanical superposition states of microtubules, which are part of the cytoskeleton in neural cells, ability to rationalize brain activities leading to

thoughts, feelings, sense of self, and transitions of consciousness; has gained much attention (Hameroff and Penrose, 2014). However, consciousness being soluble in anaesthetics such as chloroform led to question this hypothesis (Hewitt, 2014). Nevertheless, such qualitative arguments are not easy to prove through experimental procedures. Tegmark (2000) reported that quantum coherence of the ions involved in neural dynamics would be destroyed long before macroscopic dynamics could be influenced. This poses a question on how the macroscopic dynamics of neural nets arise from smaller scale coherent dynamics. Phosphorus was proposed to act as a neural qubit allowing quantum processing to occur in the brain. The quantum state of phosphorous molecules is protected through binding with calcium ions together forming special molecules known as Posner molecules (Fisher, 2015).

## Future directions and challenges

There are numerous applications of quantum biology which impact a large number of technologies including energy, environment, health, sensing, and information technologies (Adams and Petruccione, 2020; Kim et al., 2021; Marais et al., 2018). Quantum information processing in living cells can be adopted in the quantum computation industry. In this respect, there are significant developments regarding the design of artificial photosynthetic systems (Olaya-Castro et al., 2013; Scholes et al., 2012). One of the main objectives of developing artificial photosynthetic systems is converting solar energy into biofuels. The development of quantum computation through artificial neural network models that are widely used in machine learning is one of the renewing branches of science (McCulloch and Pitts, 1943). Moreover, the development of the field of quantum biology will accelerate the human attempts to establish life on other planets.

However, there are several challenges in the development of this field; including lack of measurement techniques that facilitate the understanding of the observed phenomena and certain technological bottlenecks. It is not easy to apply the currently available measurement methods for *in vivo* measurements. Furthermore, quantum biology requires both experimental and computational facilities which are expensive. In Sri Lanka, it is necessary to implement programmes to promote awareness, education and research for the active engagement to reap the benefits of quantum biology.

## Conclusion

Quantum biology combines contributions from different disciplines including physics, molecular biology, biochemistry, mathematics, engineering, chemistry and spectroscopy etc. and thus can lead to a paradigm shift in biology. Certain important advances have already been made in several biological phenomena utilizing the quantum biology approaches. However, there are numerous issues to be addressed and challenges to overcome for further advancement of the field of Biology using quantum approaches. One of the best technologies to resolve the quantum mechanical mysteries would be to utilize nanotechnology approaches to build biomimetic systems that are simpler than their biological counterparts and thereby more amenable to mathematical modelling and computation. In order to overcome most of the challenges, there is a requirement for scientists who are capable of working across disciplines of physics, chemistry and biology. Revealing of more information regarding the non-trivial role of quantum mechanics in the biological processes will help make precise explanations regarding some of nature's mysteries which will ultimately lead to the enhancement of science-based technology.

## References

Abbot, D, Davies, Paul CW and Pati AK (Eds.) 2008. Quantum Aspects of Life, Imperial College Press

Adams, B and Petruccione, F 2020. Quantum effects in the brain: A review. *AVS Quantum Science*, 2.

Agarval R, Krueger BP, Scholes GD, Yang M, Yom J, Mets L and Fleming GR. 2000. Ultrafast energy transfer in LHC-II revealed by three-pulse photon echo peak shift measurements. *The Journal of Physical Chemistry B*, 104, pp. 2908–2918.

Ball, P 2011. The dawn of quantum biology. *Nature*, 474, pp. 272-274.

Brixner T, Stenger J, Vaswani HM, Cho M, Blankenship RE and Fleming GR 2005. Two-dimensional spectroscopy of electronic couplings in photosynthesis. *Nature*, 434, pp.625–628.

Brookes JC, Hartoutsiou F, Horsfield AP and Stoneham AM 2007. Could Humans Recognize Odor by Phonon Assisted Tunneling? *Physical Review Letters*, 98, pp. 038101.

Chachisvilis M, Pullerits T, Jones MR, Hunter CN and Sundstro¨m V 1994. Vibrational dynamics in the light-harvesting complexes of the photosynthetic bacterium *Rhodobacter sphaeroides*. *Chemical Physics Letters*, 224, pp. 345–354.

Chen, GY, Lambert, N, Li, CM, Chen, YN and Nori, F 2013. Rerouting excitation transfers in the Fenna-Matthews-Olson complex. Physical Review E, pp. 88.

Collini, E, Wong, CY, Wilk, KE, Curmi, PMG, Brumer, P and Scholes, G 2010. Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature*, 463, pp. 644-647.

Davies, PCW 2003. The Origin of Life. Penguin, London (Previous title: The Fifth Miracle. Penguin, London and Simon and Schuster, New York.

Delgado M, Gorlich S, Longbotham JE, ScruttonNS, Hay S, et al. 2017. Convergence of theory and experiment on the role of preorganization, quantum tunneling and enzyme motions into flavoenzyme-catalyzed hydride transfer. *ACS Catalysis*, 7, pp. 3190-3198.

Doll KM and Finke RG 2003. A compelling experimental test of the hypothesis that enzymes have evolved to enhance quantum mechanical tunneling in hydrogen transfer reactions: The ß-neopentylcobalamin system combined with prior adocobalamin data. *Inorganic Chemistry*, 42, pp.4849-4856.

Engel GS, Calhoun TR, Read EL, Ahn TK, Mancal T, Cheng YC, Blankenship RE and Fleming GR 2007. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature*, 446, pp. 782–786.

Farhi, E and Gutmann, S 1998. Quantum computation and decision trees. *Physical Review A*, 58, pp. 915–929.

Fisher MPA 2015. Quantum cognition: the possibility of processing with nuclear spins in the brain. Annals of Physics, 362, pp. 593–602.

Fleming G et al. 2011. Quantum effects in biology. *Procedia Chemistry*, 3(1), pp.38-57.

Frick L, Neela JP and Wolfenden R 1987. Transition state stabilization by deaminases: Rates of nonenzymatic hydrolysis of adenosine and cytidine. *Bioorganic Chemistry*, 15, pp. 100-108.

Georgescu, IM, Ashhab, S and Nori, F 2014. Reviews of Modern Physics, 86, pp.153

Gerlich, S, Eibenberger S, Tomandl, M, Nimmrichter, S, Hornberger, K, Fagan, PJ, Tüxen, J, Mayor, M and Arndt, M 2011. Quantum interference of large organic molecules. *Nature*, 2, pp.263.

Gross J and Lapiere CM 1962. Collagenolytic activity in amphibian tissues: A tissue culture assay. *The Proceedings of the National Academy of Sciences USA*, 48, pp. 1014-1022.

Hameroff S and Penrose R 2014. Consciousness in the universe: a review of the 'Orch OR' theory. *Physics of Life Review*, 11, pp. 39–78.

Hewitt J 2014. Electron spin changes as a general mechanism for general anesthesia? https://phys.org/news/2014-08-electron-mechanism-anesthesia. Accessed on 28th of May 2022.

Kim, Y, Bertagna, F, D'Souza, EM, Heyes, DJ, Johannissen, LO, Nery, ET, Pantelias, A, Sanchez-Pedreno Jimenez, A, Slocombe, L and Spencer, MG 2021. Quantum Biology: An Update and Perspective. Quantum Reports, 3, pp. 1–48.

Klinman JP and Kohen A 2014. Evolutionary aspects of enzyme dynamics. *Journal of Biological Chemistry*, 289, pp. 30205-30212.

Kominis, IK 2012. Magnetic sensitivity and entanglement dynamics of the chemical compass. *Chemical Physics Letters*, 542, pp.143-146.

Konik, R 2021. Quantum coherence confined. *Nature physics*, 17, pp. 669-670.

Kumble R, Palese S, Visschers RW, Dutton PL and Hochstrasser RM 1996. Ultrafast dynamics within the B820 subunit from the core (LH-1) antenna complex. *Chemical Physics Letters*, 261, pp. 396–404.

Lambert, N, Chen, YN, Cheng, YC, Li, CH, Chen, GY and Nori, F 2013. Quantum biology. *Nature Physics*, 9, pp. 10-18.

Lloyd, S 2002. Computational capacity of the universe, *Physical Review Letters*, 88, pp.7901-7904.

Lockwood M 1989. Mind, body and the quantum. Oxford, UK: Oxford University Press.

Marais, A et al. 2018. The future of quantum biology. *Journal of Royal Society Interface*, 15.

McCulloch WS and Pitts W 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletins of Mathematical Biophysics*, 5, pp.115–133. McFadden, J 2001. Quantum Biology. Norton, New York.

McFadden, J and Al-Khalili, J 1999. A quantum mechanical model of adaptive mutations. *Biosystems*, 50, pp. 203–211.

Moser CC, Page CC and Dutton PL 2006. Darwin at the molecular scale: selection and variance in electron tunnelling proteins including cytochrome c oxidase. *Philosophical Transactions of the Royal Society B*, 361, pp. 1295-1305.

Novoderezhkin VI, Monshouwer R and van Grondelle R 2000. Electronic and vibrational coherence in the core light-harvesting antenna of Rhodopseudomonas viridis. *The Journal of Physical Chemistry B*, 104, pp. 12 056–12 071.

Novoderezhkin VI, Yakovlev AG, van Grondelle R and Shuvalov VA 2004. Coherent nuclear and electronic dynamics in primary charge separation in photosynthetic reaction centers: a Redfield theory approach. *The Journal of Physical Chemistry B*, 108, pp. 7445–7457.

Olaya-Castro A, Nazir A, Fleming GR 2013. Quantum-coherent energy transfer: implications for biology and new energy technologies. *Philosophical Transactions of the Royal Society A*, 370, pp. 3613–3617.

Panichayangangkoon, G, Hayes, D, Fransted, KA, Caram, JR, Harel, E, Wen, J, Blankenship, RE, Engel, GS 2010. Long-lived quantum coherence in photosynthetic complexes at physiological temperature, *Proceedings of the National Academy of Science*, 107, pp. 12766-12770.

Ritz, T, Yoshii, T, Helfrich-Forster, C and Ahmad, M. 2010. Cryptochrome: a photoreceptor with the properties of a magnetoreceptor? *Communicative and Integrative Biology*, 3, pp.24–27.

Romero-Isart, O, Juan, ML, Quidant, R and Cirac, JI 2010. Toward quantum superposition of living organisms. *New Journal of Physics,* 12

Sarovar, M, Ishizaki, A, Fleming, GR and Whaley, KB 2010. Quantum entanglement in photosynthetic light-harvesting complexes. *Nature Physics*, 6, pp. 462–467.

Schlau-Cohen GS, Calhoun TR, Ginsberg NS, Read EL, Ballottari M, Bassi R, van Grondelle R and Fleming GR 2009. Pathways of energy flow in LHCII from two-dimensional electronic spectroscopy. *The Journal of Physical Chemistry B*, 113(15), pp. 352–15 363.

Scholes GD, Mirkovic T, Turner DB, Fassioli F and Buchleitner A 2012. Solar light harvesting by energy transfer: from ecology to coherence. *Energy and Environmental Science Journal*, 5, pp.9374–9393.

Stapp HP 2009. Mindful universe: quantum mechanics and the participating observer. Berlin, Germany: Springer.

Tegmark M 2000. Importance of quantum decoherence in brain processes. *Physical Review E*, 61, pp. 4194–4206.

Turin L 1996. A spectroscopic mechanism for primary olfactory reception. *Chemical Senses*, 21, pp. 773–791.

Vedral, V 2011. Living in an Quantum World, *Scientific American*, pp. 38-43.

Vos M, Jones MR, Hunter CN, Breton J, Lambry JC and Martin JL 1994. Coherent nuclear dynamics at room temperature in bacterial reaction centers. *The Proceedings of the National Academy of Sciences USA*, 91, pp. 12 701–12 705.

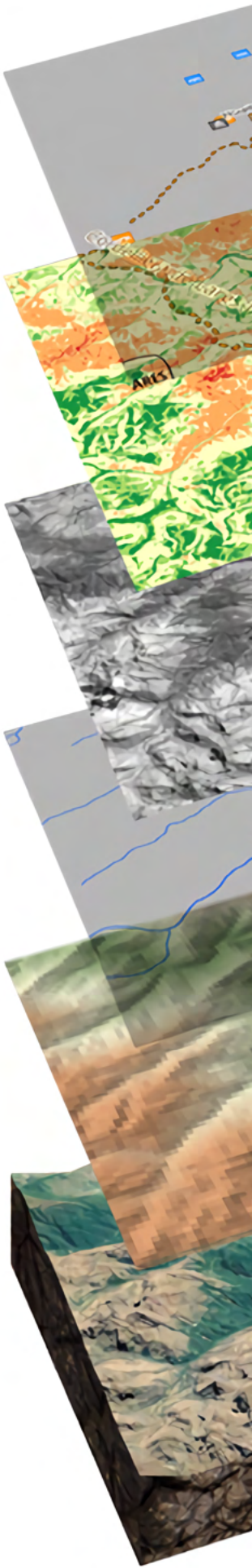
Wiltschko, W and Wiltschko, R 2005. Magnetic orientation and magnetoreception in birds and other animals. *Journal of Comparative Physiology A-Neuroethology Sensory Neural and Behavioral Physiology*, 191(8), pp. 675-693.

Wiltschko, W and Wiltschko, R 2006. Magnetoreception. *Bioessays*, 28(2), 157-168.

Zhu, XG, Long, SP and Ort, DR 2008. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Current Opinion in Biotechnology*, 19, pp. 153–159.

# Chapter 5

# Applications of Geographical Information Science (GIScience) in ecological and environmental research: Sri Lankan perspectives



A. Bellanthudawa[1], M. D. K. L. Gunathilaka[2,3], D. Halwatura[4]*

[1]Department of Agricultural Engineering, Faculty of Agriculture, University of Ruhuna, Sri Lanka
[2]School of Ecology and Environment Sciences, Yunnan University, Kunming, China
[3]Department of Geography, Faculty of Arts, University of Colombo, Sri Lanka.
[4]Department of Zoology and Environment Sciences, Faculty of Science, University of Colombo, Sri Lanka.
*Corresponding author:
devan.halwatura@zoology.cmb.ac.lk

## Abstract

Understanding ecological and environmental trends necessitates a shift from traditional methods to novel technology-driven approaches. New technological advances have taken ecological and environmental research to the next level, allowing researchers to study the relationships between organisms and their environments (ecology) and the surrounding environment (environment). Ecological and environmental studies are generally classified as intensive (repeated observation of a population to gain insight into demographic processes) or extensive (carried out over larger areas than intensive studies to obtain information on the distribution and abundance of species for conservation and management purposes), whereas environmental research is more focused on broader issues such as pollution, deforestation, and global warming. Recent advances in Geographical Information Science (GIScience) have made a significant contribution to ecological and environmental research around the world. Geographical Information Systems and Remote Sensing (GIS and RS), geo-computing, ecological modeling, and artificial intelligence and machine learning are all part of GIScience. This chapter discusses GIScience advancements in environmental and ecological research from both a global and Sri Lankan standpoint. This chapter also discusses the prospects of GIScience applications in ecological and environmental research in order to identify and fill existing research gaps as well as address current ecological and environmental issues in the Sri Lankan context.

## Keywords

*Artificial Intelligence, Ecological modelling, GIS, Remote Sensing*

## Why new technologies are required over traditional methods for ecological and environmental-related studies?

Ecology is a well-established, biocentric, and versatile field concerned with the complex and convoluted interrelationships between biotic and abiotic factors at various time and space scales (Sutherland, 2006; Forman, 2016). Over the last five decades, ecology, as well as its theories and experiments, have evolved and expanded significantly, both globally and locally. The field of "environment science" gained prominence among academics, policymakers, and decision-makers following the 1972 United Nations Conference on Human and Environment. "Ecology is the 'heart' and 'backbone' of environmental science, which emerged only in the early 1970s as an anthropocentric discipline" (Gopal, 2005).

Industrialization, along with population growth, pollution, and biodiversity loss, accelerated the advancement of ecological and environmental science research. As a result, the researchers began focusing their efforts on pollution, biodiversity loss, and global climate change. However, due to the complexity of environmental issues, such as the need for long-term data for previous years' climate change studies, research has been limited to conducting comprehensive studies. As a result, GIS has evolved into a valuable tool for managing, organizing, and visualizing geographic data that connects models to provide results from environmental and ecological research and modeling. When the domains of GIS and environmental science or GIS and ecology are combined, they cover a wide range of subjects and theoretical frameworks. In the 1980s, GIS became a more common tool for researchers because different tools were used to find and look into hard environmental problems.

Hydrology, atmospheric science, town and country planning, environmental law and conservation, risk assessments, and other topics

have seen an explosion of GIS applications and research. The literature introduces a wide range of GIS applications for environmental science and ecology. Geographic distribution, environmental and ecological processes, and patterns are all highlighted. There is a fairly broad range of environmental application areas for GIS depending on the environmental domain and the environmental issue under investigation. The use of GIS in environmental and ecological research can be classified into three categories: inventory and monitoring; analysis and modeling; and visualization and communication.

## GIScience

GIScience has emerged as a potent tool for overcoming the aforementioned constraints. It has advanced ecological and environmental research, broadening the scope of ecological and environmental research globally. GIScience encompasses Geographical Information Systems and Remote Sensing (GIS and RS), ecological modeling, and artificial intelligence. The benefit of this new GIScience is that it not only helps solve current problems in the environmental and biological sectors, but it also helps predict what might happen and lets people make decisions about how to deal with and solve future problems based on predictions and forecasts.

## Geographic Information Systems (GIS) and Remote Sensing (RS)

GIS and RS are sciences that have been heavily commercialized and used in the field of biological applications, as they are used as a tool to solve biological problems in which satellite data, maps, and aerial photographs play important roles. GIS and RS have been used all over the world to

mitigate and prepare for natural disasters, as well as to strengthen resilience and adaptation to natural disasters. The effects of hurricanes on vegetation in river basins have been studied using GIS and RS, as well as their effects on canopy level resilience, resistance, and elasticity in different vegetative land uses such as deciduous forested, evergreen forested, agricultural, and grasslands (Bellanthudawa and Chang, 2021; Bellanthudawa and Chang, 2022). Floods are regarded as one of the world's major natural disasters. Because of changing climatic and environmental factors such as precipitation, they are unpredictable. GIS and RS are important not only for identifying spatial and temporal variations in the depth of flooded water but also for assessing the extent of inundation, particularly when using light detection and ranging digital elevation models (LiDAR DEM) derived from Synthetic Aperture Radar (SAR) data (Elkhrachy, 2022; Kundu et al., 2022; Nguyen et al., 2022). Also, GIS and RS can be used to improve predictions and forecasts of flood risk landslides and associated risks for a sustainable approach to flood mitigation and preparedness (Islam et al., 2022; Munawar et al., 2022; Kalubowila et al., 2021). This helps responsible stakeholders see where there are gaps in flood management.

Wildfires are another major natural disaster that occurs around the world. However, estimating carbon losses and greenhouse gas emissions, as well as the extent of wildfires, is becoming increasingly difficult. Such assessments are aided by gridded products derived from the Terra/Aqua Moderate Resolution Imaging Spectroradiometer MODIS and Landsat (Konecny et al., 2016; Nolè et al., 2022; Sirin and Medvedeva, 2022). Remote sensing is used to understand environmental parameters such as elevation and temperature (air and land surface) in order to correlate the frequency of wildfire hotspots and the causes of them. (Ma et al., 2022)

These large-scale efforts, which are backed by research studies, help to make wildfire and land management officials more aware.

Solid waste management is a difficult task because the planning, identification of locations, and assessment of actual demand for facilities and resources necessitates proper managerial decisions due to their influence on human livelihood and quality of life, the environment, and society as a whole (Mancino et al., 2022). GIS and RS have been used to select suitable areas to locate suitable solid waste collection and management centers, particularly in developing countries (Aslam et al., 2022). To meet location and project-specific objectives, studies on the expansion of landfills and waste management centers can also be conducted using GIS and RS techniques. Remote sensing air data, drone aerial imagery, and GIS are used to keep an eye on landfill gas emissions, the effects of landfills and dumps on the environment, and to make sure that landfills are operating and being managed according to the rules (Sliusar et al., 2022).

GIS and RS have proven to be effective tools for providing long-term solutions to urban heat islands and resource management. RS has been extensively used to investigate the morphology of the urban environment and its relationship to urban heat island effects, particularly investigations into urban density and influencing features that determine block morphology (Gao et al., 2022; Moazzam et al., 2022; Xiong et al., 2022). Landsat-5 and Landsat-8 imagery, for example, is primarily used in studying strategic hotspots in urban expansion projects through land surface temperature and heterogeneity assessment (Rahman et al., 2022). Along with land surface temperatures, satellite-derived vegetation indices (e.g., Enhanced Vegetation Index, Leaf Area Index, Gross Primary

Productivity) ensure that temperatures are linked to green spaces (Ranagalage et al., 2022)

.

## Ecological modeling

In general, ecological modeling is a step-by-step process that transforms theoretical concepts into quantitative paradigms. Ecological modeling allows us to formulate dynamics and bridge the complex interactions and relationships of organisms (i.e., plants, animals, and microbes) in the ecosystem with the ecosystem's abiotic components, such as soil and water. Ecological modeling can be used to look into how the biotic and abiotic parts depend on each other and how they relate to each other, as well as to find out what the hidden effects are.

These modeling tools are used to assess ecological processes, interactions, and predictions from an ecological and environmental standpoint. Ecological models, for example, aid in the assessment of eutrophication in progress and demonstrate the level of tropic status changes in wetlands (Anagnostou et al., 2017). In comparison to other traditional methods, ecological models have been more successful in integrating anthropogenic disturbances into riparian vegetation growth and distribution. So, this help measures the health of wetland management and can be used to make decisions (You et al., 2015). Predictions and evaluations are important in agriculture and sustainable productivity, nutrient cycling, and crop productivity, and modeling is used to resolve concerns. Using agro-ecosystem simulation models to understand the biogeochemical cycles of minerals such as nitrogen and phosphorus as well as the hydrology associated with them, allows for the assessment of natural resources and capital for sustainable agriculture (Taghikhah et al., 2022). Modeling approaches are also used to investigate the functional profiles of soil carbon

and nitrogen in chemical and organic fertilizer applications. Also, ecological modeling has gotten a lot of attention in the context of soil enzyme indicators in microbial population enhancement because it affects the optimization of good soil conditions, which leads to higher crop yield and productivity (Wang et al., 2022).

Landscape ecology and risk assessment is a new field. Modeling the changes and dynamics of land use and land class classifications is important in risk assessment, and these applications can be used in endangered species conservation practices and risk quantification (Wang et al., 2020). These ecological models promote the interdependence and landscape functions of floral and faunal species while also providing detailed information on the extent of organisms' movement and distribution (Darvishi et al., 2020).

## Artificial Intelligence and Machine Learning

Artificial intelligence (AI) has received significant attention in the scientific community in recent years due to its broad applicability in a variety of fields, including biology. Water treatment and desalination is a developing industry with numerous studies being conducted to optimize the process and support long-term solutions for removing pollutants and harmful toxic substances from water and alleviating water scarcity (Alam et al., 2022; Hemanand et al., 2022). Furthermore, given the importance of water treatment and remediation, AI has been widely used in adsorption, membrane filtration, water quality parameters, and water quality index monitoring approaches, overcoming several challenges such as low efficiency and transparency in existing traditional methods as well as low model reproducibility (Lowe et al., 2022; Ramesh et al., 2022).

The breakthroughs in AI computational and research have resulted in an explosion of AI applications in the field of wildlife and biodiversity

conservation. Interestingly, scientists have used modern computer-based technology to record the echolocation calls of bat species, and this method has been useful in determining rare bat species for conservation practices (Tabak et al., 2022). Individual animal identity is another important biometric criterion in conservation. However, AI with computer-based techniques outperforms traditional DNA profiling in terms of assisting the environment in understanding individual wildlife, their movements, and the invasiveness of animal species (Congdon et al., 2022; Vidal et al., 2021). The majority of the aspects of diversity and conservation are measured by abundance and species richness. Using decision tree and random forest methods, AI and machine learning have successfully predicted the abundance and richness of fishes, spiders, and small mammals (Baltensperger and Huettmann, 2015; Smoliski and Radtke, 2017; Andek et al., 2020).

AI combined with spatial analysis has been used to predict, monitor, and evaluate global deforestation rates, assisting in the victory over deforestation's challenges. Forest managers and technicians have attempted to integrate sensors and artificial intelligence (AI) on tree canopies and tree stems to record the noise of machines such as chainsaws, which alarm the noise of cutting or felling trees over other noises (Shivaprakash et al., 2022). Furthermore, many reforestation projects in deforested areas have used AI to promote sustainable forest growth, as these activities are ultimately responsible for carbon sinks. Furthermore, we can identify potential planting areas, seed germination success, and nutrient-deficient areas using AI-based automation technology with digital intelligence (Shivaprakash et al., 2022).

## Sri Lankan perspectives of using GIS science in ecological and environmental related studies (GIS, RS, ecological modeling, artificial intelligence)

A number of issues impede the use of modern technology in developing countries like Sri Lanka. The application of GIS science in Sri Lanka is more closely related to licenses, patents, and technical issues with computer hardware and software. Consistent data collection is also lacking because there are currently no regular data collection methodologies for ground truthing. GIS and RS go far beyond "button knowledge," which cannot be grasped in a matter of weeks. As a result, due to a lack of sufficiently skilled personnel who can handle GIS and RS, the use of GIS and RS in Sri Lanka is limited to more collaborative work with people from other countries. despite the fact that local staff can be trained through short courses or through foreign training. Those who pursue foreign training may not return, which is a serious issue that must be addressed. Facilitating GIS systems could contribute significantly to the government's recent digitalization strategy, which is on the right track and will benefit Sri Lanka. Even though there have been problems, ICTA, the National Water Supply and Drainage Board, the Department of Wildlife Conservation, the Central Environmental Authority, and a number of other government institutions, including universities, have moved through several stages of working together to deploy GIS.

In September 2022, a literature survey on the evolution of research from GIS and RS to artificial intelligence and machine learning on ecology, environment, and biology in Sri Lanka was conducted using bibliometric analysis and Google Scholar as the search platform. For the analysis, journal articles, conference proceedings, and abstracts published in English between 2000 and 2021 were chosen. The terms "GIS and RS,

Ecological/Biological/Urban/Land use modeling, Artificial Intelligence" were used to search the evolutionary terminologies. Furthermore, the title words "Ecology, Biology, and Environment" were used to search for publications. The year, the title of the study, area of interest, application, and reference of all publications were recorded. The intensity of publication and its trend in the field of environment, biology, and ecology research in the Sri Lankan context was investigated.

Figure 1 depicts the evolution of literature in ecology, the environment, and biology from GIS and RS to AI and machine learning. The majority of the literature (55.7%) is based on GIS and RS, with modeling accounting for 41.8%. Furthermore, the majority of the literature (57.4%) between 2000 and 2020 focused on modeling techniques for studies related to the environment, ecology, and biology, while GIS and RS were used at a rate of 40.4%. (Figure 1) Aside from that, in the period 2020–2022, 45.7% were in GIS and RS, 34.3% in modeling, and 20% in artificial intelligence and machine learning. This difference backs up the fact that modeling, artificial intelligence, and machine learning have been studied a lot in recent years as scientists have become more interested in them.

GIS and RS were used extensively to study land use and land cover, climate change, terrestrial ecosystems, aquatic ecosystems and wetlands, natural resources (soil), health and disease, and disasters (Alahakoon et al., 2021; Rathnayake et al., 2020; Gunathilaka, 2020; Fonseka et al., 2019; Saparamadu et al., 2018; Ranagalage et al., 2018; Adikaram et al., 2017). The use of GIS and RS in ecological modeling, on the other hand, is limited. Piyasinghe et al. (2018) investigated the spatial distribution of *Austroeupatorium inulifolium* (Pathan paalu) and *Prosopis juliflora* (Kalapu andara) (Gunawardane et al., 2015). Ecological niche modeling has identified highly suitable areas for *Pseudophilautus zorro* (Gannoruwa

shrub frog) (Rupasinghe et al., 2021), which is a significant step forward in conservation biology. Amarasinghe et al. (2021) used open access GIS to investigate the niche dynamics of Memecylon sp. in response to climate change effects. Soil erosion modeling with Integrated Valuation of Ecosystem Services and Tradeoffs Another topic that has recently gotten some attention is (Udayakumara and Gunawardena, 2022). Despite the fact that there are studies that incorporate GIS and RS, it is necessary to expand the current accessibility of high-quality remotely sensed data for emerging economies. This will expand the use of GIS and RS in ecology and the environment in a more sustainable manner.

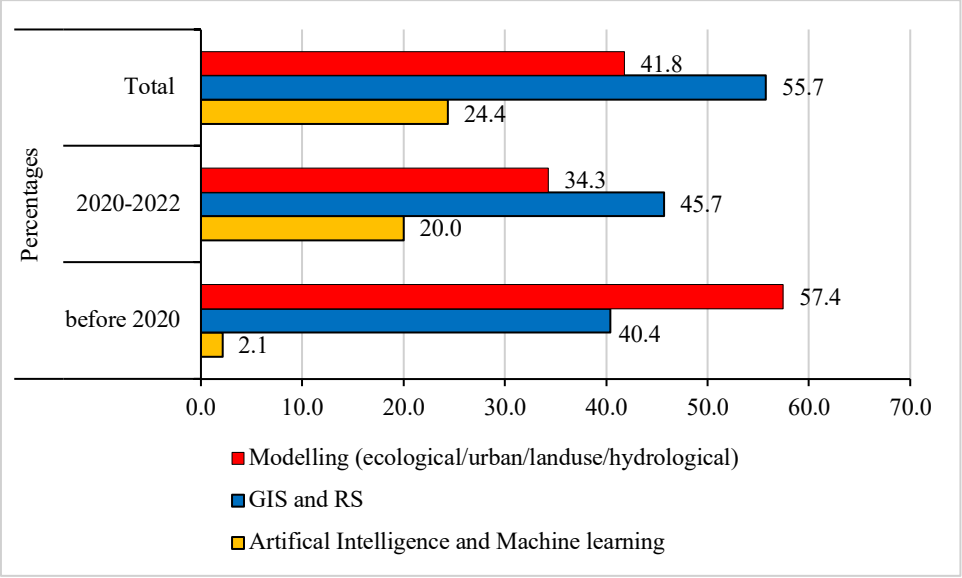



**Figure. 1.** Evolution of literature in the context of ecology, environment, and biology from the GIS and RS to Artificial intelligence and machine learning.

## Conclusion

### Moving forward: existing research gaps and addressing current ecological and environmental issues in the Sri Lankan context

There is yet much to be done in GIS science by making users able to interact with GIS and, would increase the use and applications of GIS science. However, the high-cost software and the licensing made it hard for the users to use a GIS to experience a virtual world. Building a virtual geographic scenario-based GIS and inventing geographic data-originated data analysis methodologies are the next-level goals that should be integrated with GIS. However, applications of GIS and RS in the environment and ecology in Sri Lanka are still limited to a particular frame due to the high cost of GIS and RS implementation. Therefore, compared to developed nations, environmental and ecological problems in Sri Lanka, such as the effects of climate change, natural disasters, the shrinking of wetland areas, the degradation of habitats, environmental pollution, the disposal of waste, and the depletion of resources, need to be studied in a broader way using GIS science to blend the ground data with real-time digital data. In fulfilling the broader use of GIS and RS, Sri Lanka has to undergo special strategies (Sri Lanka Spatial Data Infrastructure Strategy, 2020). Simplifying access to government spatial information is the first important strategy. Effective management of spatial data supply chains (automating information flow from top to bottom), spatial data sharing horizontally across all the relevant institutions, consumer-centric policy implementation, expand partnership across spatial data community locally and globally, and implement spatial data consumer culture are identified strategic areas envision to fulfil in the future needs.

## References

Adikaram, M., Pitawala, A., Ishiga, H., Jayawardena, D. (2017). Spatial distribution, enrichment, and source of environmentally important elements in Batticaloa lagoon, Sri Lanka. Environ Sci Pollut Res 24, 2089–2099 DOI 10.1007/s11356-016-7973-6.

Alahacoon, N.; Edirisinghe, M. (2021). Spatial Variability of Rainfall Trends in Sri Lanka from 1989 to 2019 as an Indication of Climate Change. ISPRS *Int. J. Geo-Inf.* 10, 84. https://doi.org/10.3390/ijgi10020084.

Alam, G., Ihsanullah, I., Naushad, M., & Sillanpää, M. (2022). Applications of artificial intelligence in water treatment for optimization and automation of adsorption processes: Recent advances and prospects. *Chemical Engineering Journal*, *427*, 130011.

Amarasinghe, P., Barve, N., Kathriarachchi, H., Loiselle, B., Cellinese, N., (2021) Niche dynamics of Memecylon in Sri Lanka: Distribution pattern, climate change effects, and conservation priorities. Ecology and Evolution, 11, 24. DOI 10.1002/ece3.8415.

Anagnostou, E., Gianni, A., & Zacharias, I. (2017). Ecological modeling and eutrophication—a review. *Natural Resource Modeling*, *30*(3), e12130.

Aslam, B., Maqsoom, A., Tahir, M. D., Ullah, F., Rehman, M. S. U., & Albattah, M. (2022). Identifying and Ranking Landfill Sites for Municipal Solid Waste Management: An Integrated Remote Sensing and GIS Approach. *Buildings*, *12*(5), 605.

Baltensperger, A. P., & Huettmann, F. (2015). Predictive spatial niche and biodiversity hotspot models for small mammal communities in Alaska: applying machine-learning to conservation planning. *Landscape Ecology*, *30*(4), 681-697.

Batty, M., Lin, H., and Chen, M., (2017). Virtual realities, analogies, and technologies in geography. In: B. Warf, eds. Handbook on geographies of technology. London, UK: Edward Elgar Publishing, 96–110.

Bellanthudawa, B. K. A., & Chang, N. B. (2021). Hurricane Irma's impact on biophysical and biochemical features of canopy vegetation in the Santa Fe River Basin, Florida. *International Journal of Applied Earth Observation and Geoinformation*, *102*, 102427.

Bellanthudawa, B. K. A., & Chang, N. B. (2022). Spectral index-based time series analysis of canopy resistance and resilience in a watershed under intermittent weather changes. *Ecological Informatics*, *69*, 101666.

Gopal, Brij. "Issues in Ecological Research and Environmental Management." Bulletin of the National Institute of Ecology 15 (2005): 277-286.

Breunig, M., and S. Zlatanova. (2011). 3D geo-database research: Retrospective and future directions. *Computers & Geosciences* 37:781–803. DOI: 10.1016/j.cageo.2010.04.016.

Čandek, K., Pristovšek Čandek, U., & Kuntner, M. (2020). Machine learning approaches identify male body size as the most accurate predictor of species richness. *BMC biology*, *18*(1), 1-16.

Chen, M., Yang, C., Hou, T., Lü, G., Wen, Y., Yue, S. (2018). Developing a data model for understanding geographical analysis models with consideration of their evolution and application processes. Transactions in GIS. DOI 10.1111/tgis.12484.

Chen, M., Tao, H., Lin, H., Wen, Y. (2011). A visualization method for geographic conceptual modeling. Annals of GIS, 17 (1), 15–29. DOI 10.1080/19475683.2011.558024.

Congdon, J. V., Hosseini, M., Gading, E. F., Masousi, M., Franke, M., & MacDonald, S. E. (2022). The Future of Artificial Intelligence in Monitoring Animal Identification, Health, and Behaviour. *Animals*, *12*(13), 1711.

Crooks, A.T. and Castle, C.J.E. (2011). The integration of agent-based modeling and geographical information for geospatial simulation. In: Agent-Based Models of Geographical Systems, A.J. Heppenstall, et al. (eds), 219–251. DOI 10.1177/1753193410390130.

Darvishi, A., Yousefi, M., & Marull, J. (2020). Modeling landscape ecological assessments of land use and cover change scenarios. Application to the Bojnourd Metropolitan Area (NE Iran). *Land Use Policy*, *99*, 105098.

Elkhrachy, I. (2022). Flash flood water depth estimation using SAR images, digital elevation models, and machine learning algorithms. *Remote Sensing*, *14*(3), 440.

Fonseka, H., Zhang, H., Sun, Y., Su, H., Lin, H., & Lin, Y. (2019). Urbanization and Its Impacts on Land Surface Temperature in Colombo Metropolitan Area, Sri Lanka from 1988 to 2016. Remote sensing, 11, 1-18.

Forman, R. T. (2016). Urban ecology principles are urban ecology and natural area ecology are different. *Landscape Ecology*, *31*(8), 1653-1662.

Gao, Y., Zhao, J., & Han, L. (2022). Exploring the spatial heterogeneity of the urban heat island effect and its relationship to block morphology with the geographically weighted regression model. *Sustainable Cities and Society*, *76*, 103431.

Granell, C., Havlik, D., Schade, S., Sabeur, Z., Delaney, C., Pielorz, J., Uslander, T., Mazzetti, P., Schleidt, K., Kobernus, M., Havlik, F., Bodsberg, N. R., Berre, A., Mon, J. L. (2016). Future Internet technologies for environmental applications. *Environmental Modelling & Software*, 78, 1–15. DOI 10.1016/j.envsoft.2015.12.015.

Gunathilaka, M. D. K. L. (2020) 'Detecting the future vulnerability of degradation of mangrove species through assessing the current status of the ecosystem', *International Journal of Scientific Research and Engineering Development*, vol.3 issue3, pp. 1145-1155.

Gunawardena, A., Fernando, T., Nissanka, S., Dayawansa, N. (2015). Assessment of spatial distribution and estimation of biomass of Prosopis juliflora (Sw.) DC. in Puttlam to Mannar region of Sri Lanka using remote sensing and GIS. Tropical Agricultural Research, 25(2), 228–239.

Hemanand, D., Mishra, N., Premalatha, G., Mavaluru, D., Vajpayee, A., Kushwaha, S., & Sahile, K. (2022). Applications of Intelligent Model to Analyze the Green Finance for Environmental Development in the Context of Artificial Intelligence. *Computational Intelligence and Neuroscience*, *2022*.

Hugues, O., Cieutat, J.M., and Guitton, P., 2011. GIS and augmented reality: state of the art and issues. In: Furht, ed. Handbook of augmented reality. Berlin: Springer, 721–740

Islam, M. M., Ujiie, K., Noguchi, R., & Ahamed, T. (2022). Flash flood-induced vulnerability and need assessment of wetlands using remote sensing, GIS, and econometric models. *Remote Sensing Applications: Society and Environment*, *25*, 100692.

Jerrett, Michael, Sara Gale, and Caitlin Kontgis. 2010. Spatial modeling in environmental and public health research. International *Journal of Environmental Research and Public Health* 7.4 (April): 1302–1329.DOI: 10.3390/ijerph7041302.

Kalubowila, P., Lokupitiya, E., Halwatura, D., & Jayathissa, G. (2021). Threshold rainfall ranges for landslide occurrence in Matara district of Sri Lanka and findings on community emergency preparedness. International Journal of Disaster Risk Reduction, 52, 101944.

Konecny, K., Ballhorn, U., Navratil, P., Jubanski, J., Page, S. E., Tansey, K., ... & Siegert, F. (2016). Variable carbon losses from recurrent fires in drained tropical peatlands. *Global Change Biology*, *22*(4), 1469-1480.

Kundu, S., Lakshmi, V., & Torres, R. (2022). Flood Depth Estimation during Hurricane Harvey Using Sentinel-1 and UAVSAR Data. *Remote Sensing*, *14*(6), 1450.

Li, Z.L., Yang, C., Huang, Q., Liu, K., Sun, M., Xia, J. (2017). Building models as a service to support geosciences. *Computers Environment and Urban Systems*, 61 (B), 141–152. DOI 10.1016/j.compenvurbsys.2014.06.004

Liyanage, A. S. U. (2010). Eco-geographic survey and preparation of descriptors for priority crop wild relatives of rice, vigna, banana, cinnamon, and pepper. Plant Genetic Resources Centre, Gannoruwa, Peradeniya, Sri Lanka: UNEP-GEF in situ conservation of crop wild relatives through the enhanced information management and field application project.

Lowe, M., Qin, R., & Mao, X. (2022). A review on machine learning, artificial intelligence, and smart technology in water treatment and monitoring. *Water*, *14*(9), 1384.

Lu, G.N., Yuan, L.W., and Yu, Z.Y., (2013). Challenges to development and socialization of GIS technology. Journal of Geo-Information Science, 15 (4), 483–490. DOI10.3724/SP.J.1047.2013.00483

Ma, C., Pu, R., Downs, J., & Jin, H. (2022). Characterizing Spatial Patterns of Amazon Rainforest Wildfires and Driving Factors by Using Remote Sensing and GIS Geospatial Technologies. *Geosciences*, *12*(6), 237.

Mancino, G., Console, R., Greco, M., Iacovino, C., Trivigno, M. L., & Falciano, A. (2022). Assessing Vegetation Decline Due to Pollution from Solid Waste Management by a Multitemporal Remote Sensing Approach. *Remote Sensing*, *14*(2), 428.

Moazzam, M. F. U., Doh, Y. H., & Lee, B. G. (2022). Impact of urbanization on land surface temperature and surface urban heat Island using optical remote sensing data: A case study of Jeju Island, Republic of Korea. *Building and Environment*, 109368.

Munawar, H. S., Hammad, A. W., & Waller, S. T. (2022). Remote Sensing Methods for Flood Prediction: A Review. *Sensors*, *22*(3), 960.

Nolè, A., Rita, A., Spatola, M. F., & Borghetti, M. (2022). Biogeographic variability in wildfire severity and post-fire vegetation recovery across the European forests via remote sensing-derived spectral metrics. *Science of The Total Environment*, *823*, 153807.

Nguyen, T. H., Ricci, S., Fatras, C., Piacentini, A., Delmotte, A., Lavergne, E., & Kettig, P. (2022). Improvement of Flood Extent Representation with Remote Sensing Data and Data Assimilation. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-22.

Palumbi, S. R., S. D. Gaines, H. Leslie, and R. R. Warner. (2003). New wave: High-tech tools to help marine reserve research. *Frontiers in Ecology and the Environment 1* (March): 73–79. DOI: 10.2307/3868033.

Piyasinghe, I., Gunatilake, J., Madawala, H. (2018). Mapping the distribution of invasive shrub Austroeupatorium inulifolium (Kunth) RM King & H. Rob: A case study from Sri Lanka. Ceylon Journal of Science, 47(1), 95–102.

Poplin, A., (2015). How user-friendly are online interactive maps? Survey based on experiments with heterogeneous users. *Cartography and Geographic Information Science*, 42 (4), 358–376. DOI 10.1080/15230406.2014.991427

Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, *10*(1), 3.

Ranagalage, M., Estoque, R., Zhang, X., & Murayama, Y. (2018). Spatial changes of urban heat island formation in the Colombo district, Sri Lanka: Implication for sustainability. Sustainability, 10(1367). DOI10.3390/su10051367.

Ranagalage, M., Ratnayake, S. S., Dissanayake, D. M. S. L. B., Kumar, L., Wickremasinghe, H., Vidanagama, J., ... & Muthunayake, P. (2020). Spatiotemporal variation of urban heat islands for implementing nature-based solutions: A case study of Kurunegala, Sri Lanka. *ISPRS International Journal of Geo-Information*, *9*(7), 461.

Ramesh, P., Suganya, K., Uma Maheswari, T., Paul Sebastian, S., & Sara Parwin Banu, K. (2022). Relevance of Artificial Intelligence in Wastewater Management. *The Digital Agricultural Revolution: Innovations and Challenges in Agriculture through Technology Disruptions*, 311-332.

Rathnayake, C. W.M., & Jones, S., & Benelov, M. S. (2020). Mapping LC changes over 25 years (1993-2018) in Sri Lanka using Landsat time series. *Land (Basel)*, 9(27), 2–19p.

Richards-Rissetto, H., Remondino, F., Agugiaro,G., Schwerin, J., Robertsson, J., Girardi, G. (2012). Kinect and 3D GIS in archaeology. In: Processing of 18th International Conference on Virtual Systems and Multimedia (VSMM), Milan, Italy, 331–337, 2-5 Sept.

Rupasinghe, U. D., Vandercone, R. P., & Ukuwela, K. D. (2021). Habitat associations and distribution modeling of the endangered frog *Pseudphilautus zorro* in Sri Lanka. *Herpetological Conservation and Biology*, 16(1), 38-46.

Rupasinghe, U. D., Vandercone, R. P., & Ukuwela, K. D. (2021). Habitat associations and distribution modeling of the endangered frog *Pseudphilautus zorro* in Sri Lanka. Herpetological Conservation and Biology, 16(1), 38-46.

Saparamadu, S., Yi, Z., & Zongping, Z. (2018). Temporal changes of land cover and environmental impacts; A case study in Colombo, Sri Lanka. *International Journal of Earth Environment Science*, 3(150).

Shen, S., Gong, J., Liang, J., Li, W., Zhang, D., Huang, L., Zhang, G. (2018). A heterogeneous distributed virtual geographic environment—potential application in spatiotemporal behavior experiments. *ISPRS International Journal of GeoInformation*, 7 (2), 54. DOI 10.3390/ijgi7020054

Shivaprakash, K. N., Swami, N., Mysorekar, S., Arora, R., Gangadharan, A., Vohra, K., ... & Kiesecker, J. M. (2022). Potential for Artificial Intelligence (AI) and Machine Learning (ML) Applications in Biodiversity Conservation, Managing Forests, and Related Services in India. *Sustainability*, *14*(12), 7154.

Sirin, A., & Medvedeva, M. (2022). Remote Sensing Mapping of Peat-Fire-Burnt Areas: Identification among Other Wildfires. *Remote Sensing*, *14*(1), 194.

Sirin, A., & Medvedeva, M. (2022). Remote Sensing Mapping of Peat-Fire-Burnt Areas: Identification among Other Wildfires. *Remote Sensing*, *14*(1), 194.

Sliusar, N., Filkin, T., Huber-Humer, M., & Ritzkowski, M. (2022). Drone technology in municipal solid waste management and landfill: A comprehensive review. *Waste Management*, *139*, 1-16.

Smoliński, S., & Radtke, K. (2017). Spatial prediction of demersal fish diversity in the Baltic Sea: comparison of machine learning and regression-based techniques. *ICES Journal of Marine Science*, *74*(1), 102-111.

Sri Lanka Spatial Data Infrastructure Strategy (2020). *Powering Decision Making and Innovation Using Spatial Information Technologies*.

Stockwell, C. E., Bela, M. M., Coggon, M. M., Gkatzelis, G. I., Wiggins, E., Gargulinski, E. M., ... & Warneke, C. (2022). Airborne Emission Rate Measurements Validate Remote Sensing Observations and Emission Inventories of Western US Wildfires. *Environmental Science & Technology*.

Sui, D. (2014). Opportunities and impediments for open GIS. *Transactions in GIS* 18.1 (February): 1–24. DOI: 10.1111/tgis.12075

Sutherland, W. J. (Ed.). (2006). *Ecological census techniques: a handbook*. Cambridge university press.

Tabak, M. A., Murray, K. L., Reed, A. M., Lombardi, J. A., & Bay, K. J. (2022). Automated classification of bat echolocation call recordings with artificial intelligence. *Ecological Informatics*, *68*, 101526.

Taghikhah, F., Borevitz, J., Costanza, R., & Voinov, A. (2022). design: A dynamic agroecosystem simulation model for natural capital assessment. *Ecological Modelling*, *468*, 109930.

Udayakumara, E.P.N., Gunawardena, P. Modelling soil erosion and hydropower linkages of Rantambe reservoir, Sri Lanka: towards payments for ecosystem services. Model. Earth Syst. Environ. 8, 1617–1634 (2022). DOI10.1007/s40808-021-01169-6.

Vidal, M., Wolf, N., Rosenberg, B., Harris, B. P., & Mathis, A. (2021). Perspectives on individual animal identification from biology and computer vision. *Integrative and Comparative Biology*, *61*(3), 900-916.

Voinov, A., Çöltekin, A., Chen, M., Beydoun, G. (2018). Virtual geographic environments in socio-environmental modeling: a fancy distraction or a key to communication? *International Journal of Digital Earth*, 11 (4), 408–419. DOI 10.1080/17538947.2017.1365961.

Wang, G., Gao, Q., Yang, Y., Hobbie, S. E., Reich, P. B., & Zhou, J. (2022). Soil enzymes as indicators of soil function: A step toward greater realism in microbial ecological modeling. *Global change biology*, *28*(5), 1935-1950.

Wang, G., Wang, C., Guo, Z., Dai, L., Wu, Y., Liu, H., ... & Xue, F. (2020). Integrating Maxent model and landscape ecology theory for studying spatiotemporal dynamics of habitat: Suggestions for conservation of endangered Red-crowned crane. *Ecological Indicators*, *116*, 106472.

Wen, Y.N., Chen, M., Yue, S., Zheng, P., Peng, G., Lu, G. (2016). A model-service deployment strategy for collaboratively sharing analysis models in an open web environment. *International Journal of Digital Earth*. DOI 10.1080/17538947.2015.1131340.

Wu, M.G., (2015). A spatial distribution pattern-driven spatial index. *Acta Geodaetica Et Cartographica Sinica*, 44 (1), 108–115.

Xiong, L., Li, S., Zou, B., Peng, F., Fang, X., & Xue, Y. (2022). Long Time-Series Urban Heat Island Monitoring and Driving Factors Analysis Using Remote Sensing and Geodetector. *Frontiers in Environmental Science*, *9*, 828230.

You, X., Liu, J., & Zhang, L. (2015). Ecological modeling of riparian vegetation under disturbances: A review. *Ecological modeling*, *318*, 293-300.

ISBN 978-624-5998-02-9

9 786245 998029

Thematic Publication 2022

The Institute of Biology Sri Lanka