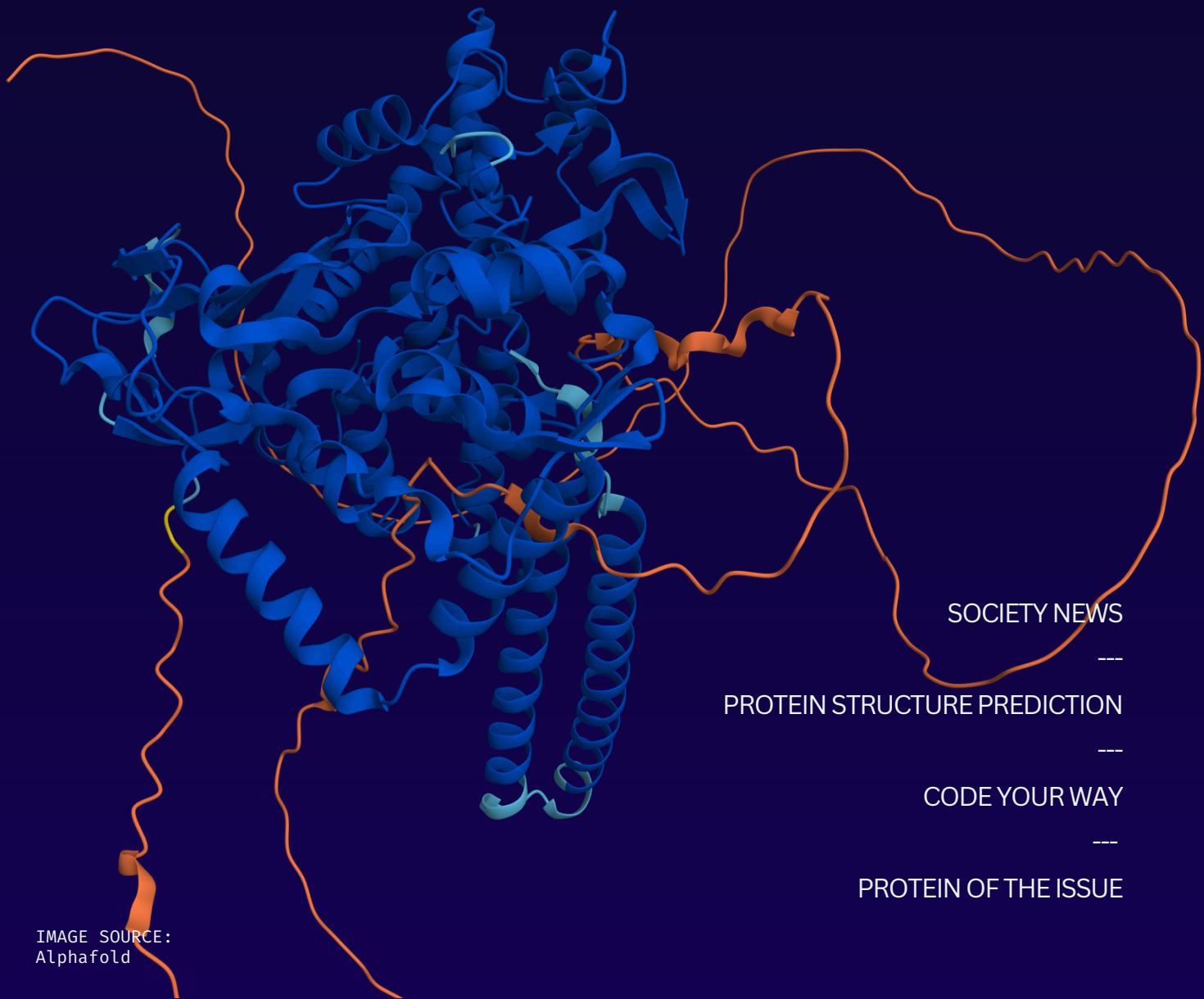


Protein Structure Prediction

Cracking the code:
predicting protein structures with bioinformatics



Society News



BINF
SOC

Recent Events.

- Bioinformatics Conference Term 3 Week 9 2024
 - Meet your Cohort Term 1 Week 2 2025
 - Movie Night Term 1 Week 6 2025
 - Lab tour Term 1 Week 7 2025
-

Upcoming Events.

- Red Bull Wings4Life Term 1 Week 11 2025
- Industry Mentoring Term 2 Weeks 1-9 2025

Combine

2024

Bioinformatics Student Symposium



BINFSOC



BININFORMATICS CONFERENCE AT USYD

In Term 3, Week 9 of 2024, 40 students from the UNSW Bioinformatics Society had the exciting opportunity to attend the COMBINE 2024 Bioinformatics Student Symposium, hosted at the University of Sydney. The symposium brought together students, researchers, and academics from across the country, making it a truly vibrant and inspiring event for anyone passionate about bioinformatics and computational biology.

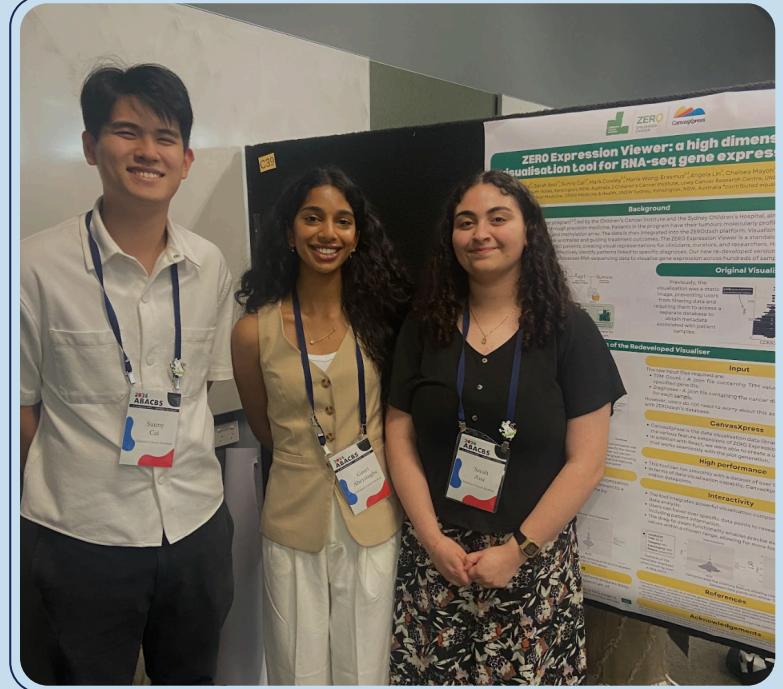
Throughout the day, our members were immersed in a diverse range of talks and poster sessions that covered the latest advancements in the field. Topics included the analysis of protein structures, integration of multi-omics data, the transformative role of artificial intelligence in biological research, and the use of computational tools to better understand disease mechanisms etc.

For many of our students, it was a chance to engage with complex biological questions through the lens of computation and data science — and to see how their own studies connect with real-world research.

The event featured presentations from Honours and PhD students, as well as keynote sessions delivered by prominent researchers from Australia's leading institutions, such as the Garvan Institute of Medical Research, Children's Cancer Institute, WEHI (Walter and Eliza Hall Institute of Medical Research), Victor Chang Cardiac Research Institute, CSIRO, Black Ochre Data Labs, and a range of universities.

It was both inspiring and eye-opening to hear about the breadth of projects underway across the country — from structural bioinformatics and cancer genomics to systems biology and AI-driven discovery.

A standout moment for our society was seeing several of our very own members present their research in the poster session. Congratulations to Anish Sanghvi, Gauri Abeysinghe, Sunny Cai, and Sarah Assi, who confidently showcased the work they had completed during their summer research programs and their DESN2000 projects. It was a proud moment to witness their hard work and dedication being recognised and appreciated by the broader bioinformatics community.



Beyond the academic sessions, the symposium also offered our students the chance to network with like-minded peers, ask questions about potential Honours and postgraduate research paths, and gain insights into the current and future directions of the field.

The atmosphere was welcoming, and the event offered a great balance between learning, professional development, and community building.

We are also incredibly grateful to UNSW Computer Science and Engineering for sponsoring our members to attend. Their support made it possible for our students to participate in this enriching experience and bring back valuable insights to our society and broader academic community.

Events like this help build a strong, connected community of future bioinformaticians, and we're excited to see even more participation in the years to come.



TESTIMONIALS



“

Through this conference, I rekindled a childhood interest in viruses and had the chance to connect with passionate individuals and their inspiring professors who specialise in this field. All to say, this experience has provided me with a clearer vision of my future.

> Xianzhi Yi

I was fascinated to learn about the issues facing sustainability in professional bioinformatics circles. The problems with getting funding to maintain the tools created in academia and keeping them consistently updated, moving away from organisations always following protocol towards always having a self-checking/evaluation of all processes.

It was also great to hear from current professionals in the careers panel, learning what their main advice is for going into the field of bioinformatics.

> Madeline Robinson

It was very moving to hear from such passionate presenters talk about their discoveries in science. It made me excited about the future of bioinformatics and after the conference I felt extremely motivated to work in this thriving field. It also made me think about what research topic I want to study in my honours project next year.

> Jaime Taitz

”



MEET YOUR COHORT



Bioinformatics Society hosted our Meet Your Cohort event, bringing together new and returning members for an evening of introductions, conversation, and community-building. Students had the opportunity to connect It was a fantastic opportunity for students across different years to connect over shared interests in bioinformatics, exchange advice, and get a feel for the supportive network within our society. Thanks to everyone who came along and helped make the event such a success—we're looking forward to an exciting year ahead!





B MOVIE NIGHT

On Thursday, March 27th, BINFSOC hosted a movie night at the Colombo building, bringing students together for an evening of entertainment and relaxation. The event provided a perfect opportunity for students to take a break from their studies and unwind after the demanding midterm period. With food and drinks provided, attendees enjoyed watching Thor: Ragnarok and unwinding after a stressful midterm period. Throughout the night, students had the chance to bond and connect in a fun, laid-back environment. The event saw a fantastic turnout, and we look forward to organising more events that bring together the bioinformatics community in the future.

B LAB TOUR

On Tuesday, April 1st, the students of BINFSOC were given the amazing opportunity to tour through Sydney Local Health District's Institute of Precision Medicine and Bioinformatics. Students explored state-of-the-art facilities and saw first-hand evidence of the groundbreaking work being done in genomics, data-driven diagnostics, and personalised healthcare solution. From seeing high-throughput sequencing technologies in action to discussing real-world applications of bioinformatics in clinical settings, it was a great experience for students to see how research is directly improving patient care and shaping the future of medicine. Thank you to the team at @sydneylocalhealth district for sharing their passion for innovation with us!



Decoding Proteins: The Bioinformatics Breakthrough

Authors: Riva Manter & Ethan Morritt



An introduction to protein structure prediction

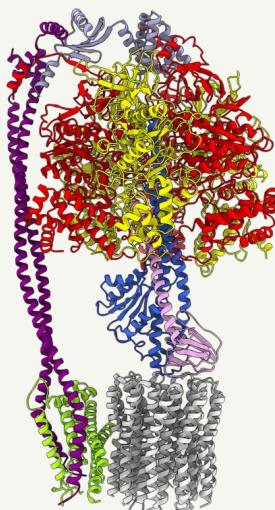
Proteins are large biomolecules that perform essential functions in the body, including roles in growth, structural support, and regulation. A singular protein is made up of chains of amino acids, folded into specific 3-D structures through multiple stages that determine their function. These structures are shaped by various chemical interactions originating from preliminary structures and can be experimentally determined through technologies such as X-ray crystallography, and cryogenic electron microscopy. However, these experimental approaches are time-consuming, labour-intensive, and costly. This is where bioinformatic methods for protein structure prediction offer a powerful alternative.

Computational methods for protein structure prediction have traditionally followed two main pathways: modelling physical interactions or analysing evolutionary history.

Analysing Physical Interactions

Physical interaction-based approaches combine our understanding of molecular driving forces with thermodynamic or kinetic simulations of protein physics (or statistical approximations of these interac-

tions). While conceptually grounded in physics, these simulations become increasingly difficult with larger proteins, where the complexity of interactions and structural dependencies make accurate modelling challenging. Along with this, protein stability is highly context-dependent due to a complex interplay of various different factors. This network of interactions within a folded protein presents further complicates accurate simulation.



Protein structure prediction of *Bacillus PS3* ATP synthase.
Image: [eLife](#)

Analysis of Evolutionary History

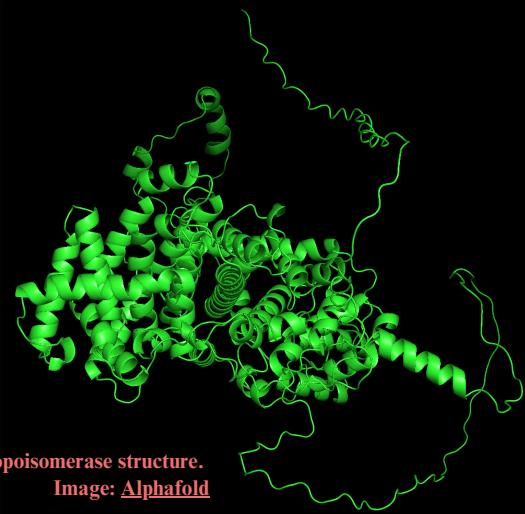
The second pathway involves analysing the evolutionary history of proteins, specifically through examining statistical relationships in the evolutionary changes of proteins. If two residues consistently change together across evolutionary lineages, it suggests they may interact structurally.

This approach has become more relevant recently due to a number of key advancements: a major growth of experimental protein structures in the Protein Data Bank (PDB), as well as breakthroughs in deep learning techniques.

AlphaFold

A notable example of this is the development of AlphaFold, a neural network-based model which has demonstrated unprecedented accuracy in the prediction of protein structures, even for proteins with no known similar structures. Behind this major breakthrough is a novel machine learning approach which combines both physical and biological insights of protein structure. AlphaFold takes the protein's amino acid sequence and aligned sequences of homologous proteins (multiple sequence alignment) as inputs in order to output a highly accurate protein structure with atomic-level precision.

To further boost accuracy, AlphaFold employs a training strategy known as self-distillation. After initial training on known PDB structures via supervised learning, the model generates its own predictions, which are then combined with the PDB data to train a second-generation network. This iterative process significantly enhances performance.



Conclusion

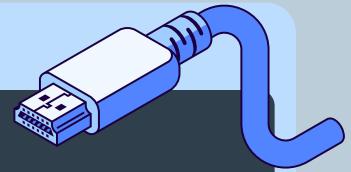
Emergence of deep learning based models such as AlphaFold mark a major shift towards algorithmic modelling in the field of protein structure prediction. By relying on neural networks to learn the relationships between sequence and structure directly from existing data, these models reduced the need for explicit statistical models or physical simulations. However, there is still room for improvement, particularly for proteins with few homologous sequences, or those whose structures are strongly influenced by interactions with other molecules. Future research is needed in these areas, as well as in interpreting the behaviour of trained neural networks, which hold significant potential for advancing our understanding of the fundamental principles underlying protein folding.

References and further reading:

- <https://www.nature.com/articles/s41586-021-03819-2>
- <https://pmc.ncbi.nlm.nih.gov/articles/PMC10928435/>
- <https://alphafold.ebi.ac.uk/>
- <https://elifesciences.org/articles/43128#content>

Code Your Way

```
$ chmod +x print_info.py  
$ ./print_info.py
```



Welcome to...

<< CODE YOUR WAY! >>

Today, we are diving into the world of Shell! A shell is a command interpreter and acts as an interface between the user and the computer! Shell is very powerful as it allows a user to store commands for repetitive tasks. It also allows us to sift and manipulate through huge sets of data.

Bioinformaticians use Shell every day to simplify their data and reach conclusions in a simple way!



ATTGCCTTAT**SHELL**
ACTGATAGCC**TIME**_

Is there a coding concept you'd love to see featured in **Code Your Way**? Let us know by emailing binfsights@binfsoc.com — we'd love to hear from you!

Basic Shell

Let's start off with some basic Shell commands!

Navigating the file system:

Several commands can be used to navigate your computer's file system!

- Show current working directory:

```
$ pwd
```

- Show what files/directories exist in our current working directory:

```
$ ls
```

- Change directory:

```
$ cd [directory]
```

- Move up one directory level:

```
$ cd ..
```

Manipulating files and directories:

We can also create, copy, and delete files and directories using the following commands!

- Create directory:

```
$ mkdir [new directory/path]
```

- Create text file:

```
$ nano [new text file name]
```

- Move a file to a different location:

```
$ mv [source path] [destination path]
```

- Copy a file/directory:

```
$ cp [file to copy] [new file]
```

- Remove a file/directory (be careful!):

```
$ rm [file/directory path]
```

Additional commands

(here are a couple of handy extras to boost your Shell skills!):

- The lines of a text file are able to be sorted using the following command:

```
$ sort [text file]
```

- We can also use pipes (the vertical bar '|') to capture output from one command and use it as input for a second command. In the example below, we are telling the shell to give us the files and directories in the current directory in alphabetical order:

```
$ ls | sort
```

More Shell

Let's do a tutorial with Shell commands!

What is grep?

grep is a useful search command that matches regular expressions (REGEX).

```
$ grep -E 'hello' file.txt
```

What does this code do? From file.txt, search for all the lines that contain 'hello'. (This is similar to Ctrl+f functionality!)

Let's say we have a list of how many birds were sighted on a given day:

```
date|bird|number  
01/02/2017|ibis|5  
03/11/2013|cockatoo|13  
24/09/2019|ibis|2900  
12/02/2020|king parrot|5  
18/06/2015|crow|14  
14/05/2011|rozella|2  
09/03/2013|rozella|5  
19/04/2018|ibis|12  
12/12/2019|crow|29  
02/07/2017|rozella|3
```

Read from Step 1) to see how you can apply all of these commands to return different data in this file!

What is cut?

tr replaces a character sequence with another.

```
$ tr A-Z a-z
```

What does this code do? Replace all uppercase with lowercase.

What is cut?

cut can be used to isolate given data by separating by a character. This character can be referred to as a delimiter.

```
$ cut -d'|' -f2
```

What does this code do? Define columns separated by the delimiter '|', and return the data in the 2nd column.

1) Let's return all lines that contain 'rozella'.

```
$ grep -E 'rozella' birds.txt
```

```
14/05/2011|rozella|2  
09/03/2013|rozella|5  
02/07/2017|rozella|3
```

2) Let's get all the birds spotted in the year 2019.

```
$ grep -E '2019' birds.txt
```

```
24/09/2019|ibis|2900  
12/12/2019|crow|29
```

Now, let's use cut and a PIPE to return only the BIRD field from all the birds spotted in 2019:

```
$ grep -E '2019' birds.txt | cut -d'|' -f2
```

```
ibis  
crow
```

3) There's a mistake on the line where there are 2900 ibis. Let's fix this by replacing 2900 with 29.

```
$ grep -E 'ibis' birds.txt | tr '2900' '29'
```

```
01/02/2017|ibis|5  
24/09/2019|ibis|29  
19/04/2018|ibis|12
```

before

```
24/09/2019|ibis|2900
```

after

```
24/09/2019|ibis|29
```

Shell: a cheatsheet!

We've put together a useful cheatsheet featuring all of the shell commands and their functions listed in this issue (plus extras!). Whether you're just starting out or simply need a quick refresher, we hope you find it helpful.

> **pwd**

Prints the full (absolute) path of the current working directory.

> **ls**

List directory contents.

> **cd [directory]**

Change directory.

> **cd ..**

Go up one directory.

> **mkdir [new directory]**

Make a new directory.

> **nano [new text file]**

Open a file in the nano text editor. If the file doesn't exist, it will be created.

> **mv [source] [destination]**

Move or rename files and directories.

> **cp [file to copy] [new file]**

Copy files and directories.

> **rm [file/directory path]**

Remove files or directories.

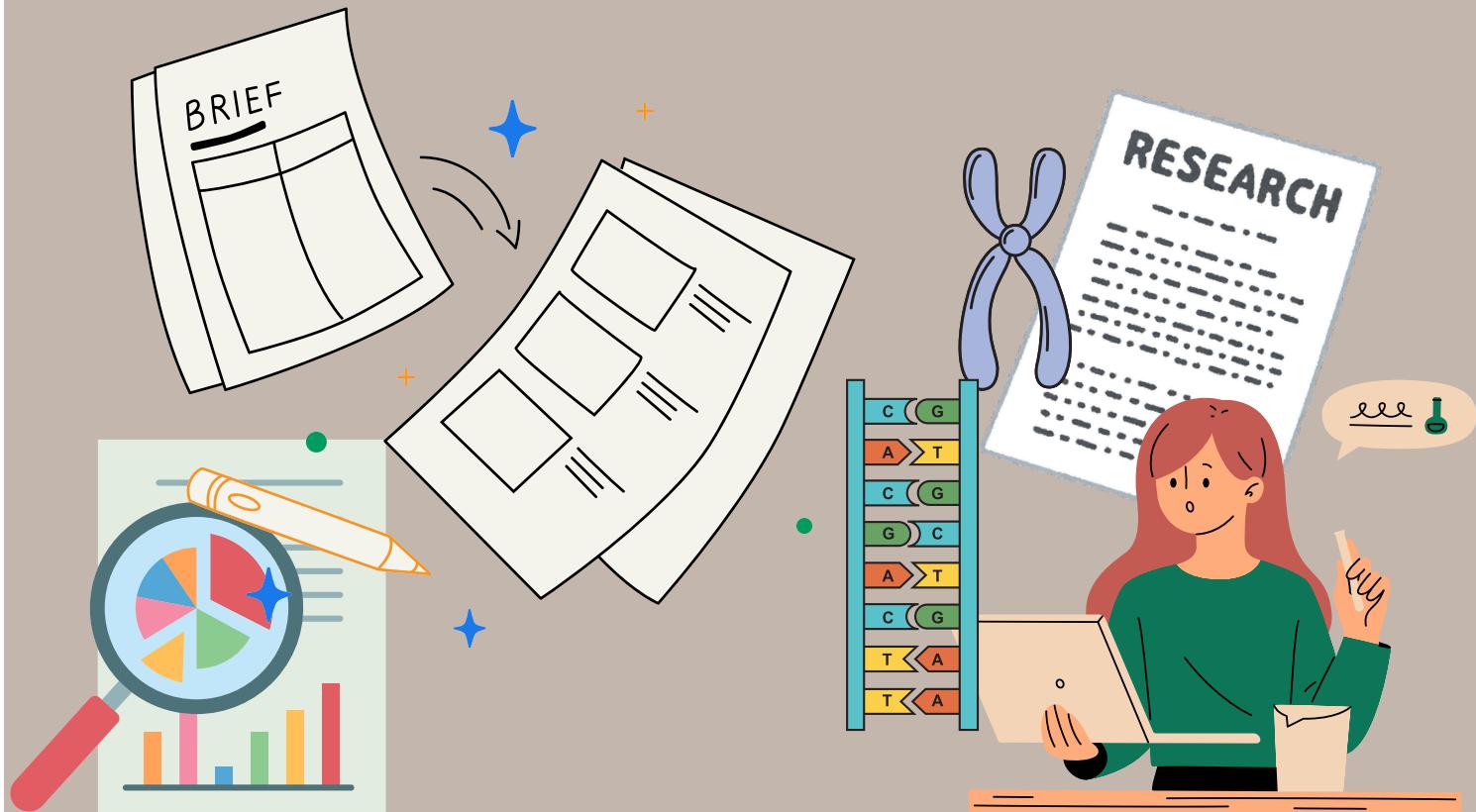


ATTCGCCTTAT**SHELL**
ACTGATAGCC**TIME**_

ASK A PhD STUDENT

Welcome to “Ask a PhD student”:

our feature where we explore the paths, challenges, and insights from those who've taken the leap into postgraduate research. In this issue, the BINFSOC team spoke with a current UNSW PhD student to gather their advice. Whether you're seriously considering a PhD or just curious about what it involves, this piece offers valuable guidance and real experiences to help you make informed decisions about your future in academia.





KRISTINA SANTUCCI

PHD CANDIDATE AT UNSW

> What first sparked your interest in bioinformatics?

During my undergraduate degree in Medical Science, I had my first genetics course and absolutely enjoyed every minute of it. From there I knew I wanted to major in genetics and decided to pursue a career in research. In my third year, I completed a course called the Biomolecular Science Laboratory Project where you undertake a term-long research project under the supervision of an academic staff member. This project was in a bioinformatics lab under the supervision of Dr Michael Janitz. It was such an amazing experience that naturally, I decided to complete an Honours degree and PhD in the same lab.

> Can you tell us a bit about your PhD research at UNSW?

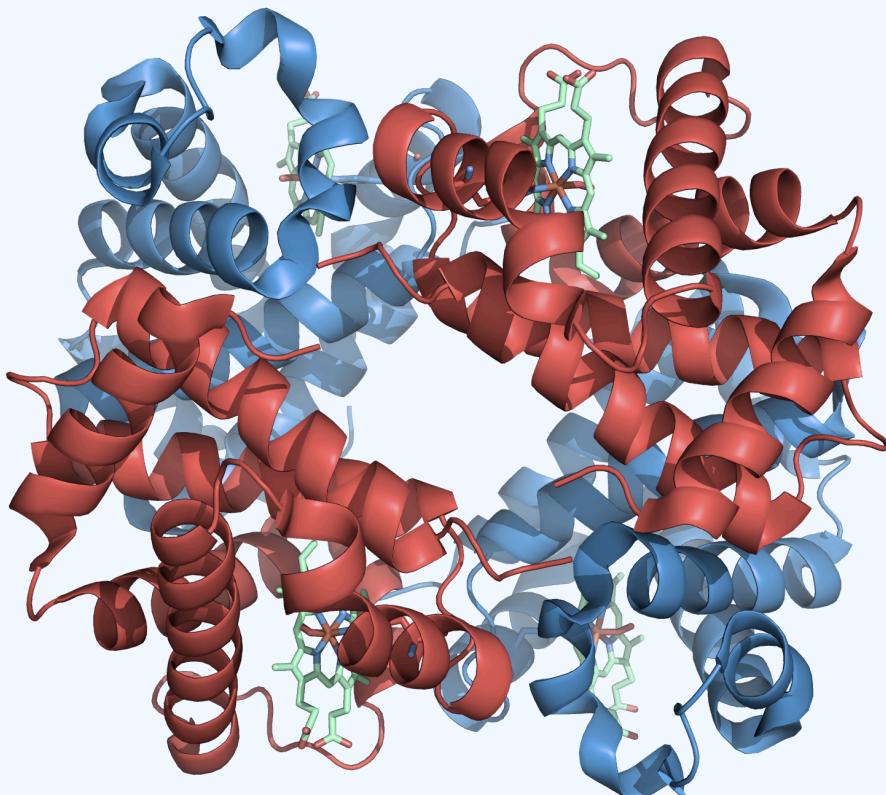
My research focuses on the discovery and analysis of new genes, RNA isoforms, and proteins through bioinformatic analysis of RNA sequencing data. So far, my experience as a PhD candidate at UNSW has been wonderful, as my supervisor and lab team are very supportive. The Graduate Research School also provides lots of training and resources to support new candidates like myself, which has been very helpful.

> Can you tell us a bit about your PhD research at UNSW?

If you are coming from a computer science background, ensure you have a strong understanding of biology and genetics. If you are coming from biology or medical science like myself, enrol in bioinformatics courses as your electives to gain skills in coding and scripting. For those in a bioinformatics degree who get the best of both worlds, work on your problem-solving skills as lots of issues that you encounter during your research haven't been encountered before. Finally, be curious, ask questions, and read! I can't stress enough how much staying on top of new developments in your field of research can assist in shaping your project and fine-tune your own knowledge in that area.

PROTEIN OF THE ISSUE

Haemoglobin



The protein responsible for **transporting oxygen** around your body through red-blood cells.

- 4 subunits, each with a **polypeptide chain (globin)**, '**Heme**' group, and **iron**
 - Oxygen binds to each iron
 - Iron content in the Haemoglobin is the reason why your blood cells are **red!**
- Each of your blood cells carry around **260 million** Haemoglobin molecules!

Our Sponsors



BINFSOC

Thank you to all our partners and sponsors
for their continuous support.



Red Bull is the number-one energy drink manufacturer, partnering with leading athletes and organisations around the world

Sponsor Us

If you would like to become a sponsor of BINFSOC,
please reach out to sponsorships@unswbinfsoc.com

Contact us



IF YOU HAVE ANY COMMENTS or feedback regarding BINFsights, please write to us at binfo@unswbinfsoc.com

We also encourage anyone to share with us anything you'd like us to take a look at, be it a bioinformatics tool that you have made or find useful; or news in the bioinformatics world that you'd like to see written about in future issues.



TO VIEW PAST AND PRESENT issues of BINFsights, check out our website at unswbinfsoc.com/binfo
Stay tuned on our Facebook page for updates regarding events and society news.

-- The BINFSOC Team

Acknowledgements



HAFSA FAHAD

PUBLICATIONS/IT EXECUTIVE & TEAM LEAD



YVONNE HUANG

VICE PRESIDENT



RIVA MANTER

PUBLICATIONS/IT DIRECTOR



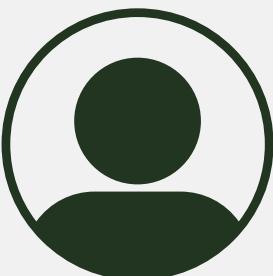
PHOEBE TANDJIRIA

PUBLICATIONS/IT DIRECTOR



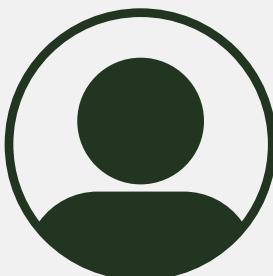
Angelina Tran

PUBLICATIONS/IT
SUBCOMMITTEE
MEMBER



Suhani Jones

PUBLICATIONS/IT
SUBCOMMITTEE
MEMBER



Ethan Morritt

PUBLICATIONS/IT
SUBCOMMITTEE
MEMBER



Zoe Brookes

PUBLICATIONS/IT
SUBCOMMITTEE
MEMBER