DAVID ZHANG

Experienced bioinformatician who enjoys learning and applying the best practices for method development. I focus on developing robust, user-friendly tools that harness transcriptomic data to improve the rate of genetic diagnosis.

🎓 EDUCATION

Present | 2017

PhD, Bioinformatics
University College London
📍 London, UK

- **Thesis**: Using transcriptomics to improve the genetic diagnosis rate of rare disease patients.

2016 | 2015

MSc, Neuroscience
University College London
📍 London, UK

- **Thesis**: The role of mitochondrial dysfunction in Xerodoma pigmentosum
- **Grade**: Merit (68%)
- Awarded post-graduate support scheme bursary (£10,000)

2015 | 2012

BSc, Biomedical science
University College London
📍 London, UK

- **Thesis**: Investigating the function of CYFIP1 in the development of rat hippocampal neurons.
- **Grade**: 2:1 (69%)

2012 | 2007

H.S.
Queen Elizabeth's School
📍 Barnet, UK

- **Grades**: Maths (A*), Biology (A*), Chemistry (A*), Sociology (A).

💼 WORK EXPERIENCE

2021

Bioinformatician internship (2 months)
Verge Genomics
📍 Remote, US

- **Goal**: Set up an aberrant splicing detection pipeline for drug target discovery in C9orf72 ALS patients.
- Used **docker** to setup a reproducible workflow[1] for running aberrant splicing analyses on an **AWS** instance.

CONTACT

- ✉ dyzhang32@gmail.com
- GitHub
- in LinkedIn
- ResearchGate

LANGUAGE SKILLS

R
Python
Git/GitHub
Bash
docker

📈 SOFTWARE & PROGRAMMING

Present
|
2020

● **Bioconductor packages**

· dasper[2]: Detection of aberrant splicing events in RNA-sequencing. **Author** and **maintainer**.

· megadepth[3]: BigWig and BAM related utilities. An R wrapper for the megadepth software developed by Chris Wilks. **Co-author** and **maintainer**.

· ODER[4]: Optimising the definition of Expressed Regions. Submitted to Bioconductor. **Co-author** and **maintainer**.

Present
|
2021

● **Python packages**

· codino[5] converts a codon design to the expected amino acid frequencies, and vice versa. **Author** and **maintainer**.

2021

● **Web scraping**

· Applied the python packages **Beautiful Soup** and **Selenium** to web scrape[6] information on all UK biotechnology companies.

2021

● **Data science blog**

· Chess-related blog post[7] was selected for the hand-on-tutorials column in Towards Data Science, which displays pieces that highlight best practices of data science.

· Applied popular data science packages in **python** to analyse[8] chess.com data.

🔗 LINKS

1: https://github.com/dzhang32/auto_splice
2: https://bioconductor.org/packages/release/bioc/html/dasper.html
3: https://bioconductor.org/packages/release/bioc/html/megadepth.html
4: https://github.com/eolagbaju/ODER
5: https://github.com/dzhang32/codino
6: https://github.com/dzhang32/biotech_web_scrape
7: https://towardsdatascience.com/how-has-the-queens-gambit-impacted-the
   -popularity-of-online-chess-43594efe5a98
8: https://github.com/dzhang32/chess