DAVID ZHANG

Experienced bioinformatician who has experience developing robust, user-friendly software in python or R to facilitate diagnostics and drug discovery.

View this CV online with links at *dzhang32.github.io/cv/*

WORK EXPERIENCE

present | 2022

Machine learning engineer

Ladder therapeutics — London, UK (remote)

- **Goal**: Implementing and engineering production-ready software leveraging RNA biology and chemistry to accelerate drug discovery.

2021

Bioinformatician internship (2 months)

Verge Genomics — London, UK (remote)

- **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.

EDUCATION

2022 | 2017

PhD, Bioinformatics

University College London — London, UK

- **Thesis**: Using transcriptomics to improve the genetic diagnosis rate of rare disease patients.
- Developed and released software that facilitate transcriptomic analyses with a focus on diagnostics.

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%)

CONTACT

- dyzhang32@gmail.com
- GitHub
- LinkedIn
- Google Scholar

LANGUAGE SKILLS

- Python
- R
- Git/GitHub
- Bash
- docker
- SQL

The long-form version of my CV with a list of selected publications is available here.

*Made with the R package **datadrivencv** and **pagedown**.*

The source code is available at github.com/dzhang32/cv.

Last updated on 2022-05-25.

2012 | 2007

H.S.

Queen Elizabeth's School — Barnet, UK

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

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 | 2022

R packages

- ggtranscript[5]: Visualising transcript structure and annotation using ggplot2. **Author** and **maintainer**.
- autorecipes[6]: Automate your recipe planning. **Author** and **maintainer**.

Present | 2021

Python packages

- autogroceries[7]: Automate your grocery shop. **Author** and **maintainer**.
- codino[8] 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[9] information on all UK biotechnology companies.

2021

Data science blog

- Chess-related blog post[10] 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[11] 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/ggtranscript
6: https://github.com/dzhang32/autorecipes
7: https://github.com/dzhang32/autogroceries
8: https://github.com/dzhang32/codino
9: https://github.com/dzhang32/biotech_web_scrape
10: https://towardsdatascience.com/how-has-the-queens-gambit-impacted-the-popularity-of-online-chess-43594efe5a98
11: https://github.com/dzhang32/chess