



CompBio Careers

William Haese-Hill

Research Software Engineer

28/11/2025

MY BACKGROUND

PhD in Mathematical Physics

2015



Loughborough University



Loughborough University Science and Enterprise Park

Software Developer
in Industry
(~7 years)



Nigel Cox / Holborn Town Hall



University
of Glasgow

Research
Software Engineer
From 2021

Maths PhD

2011-2015

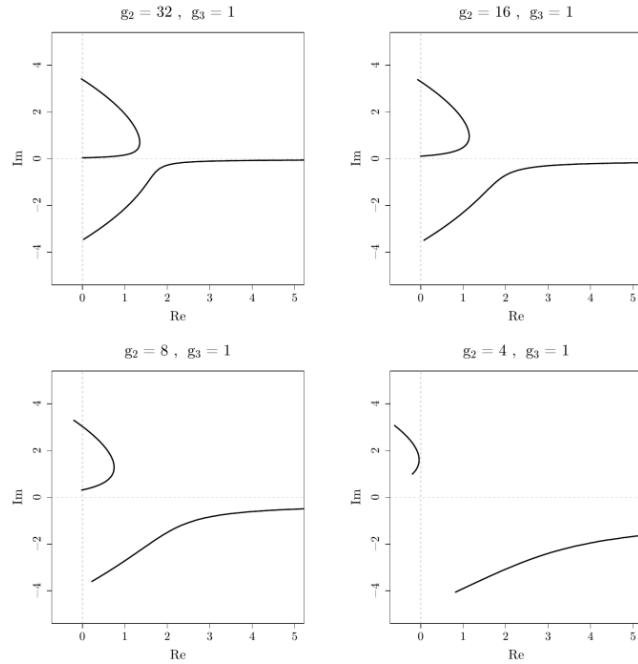
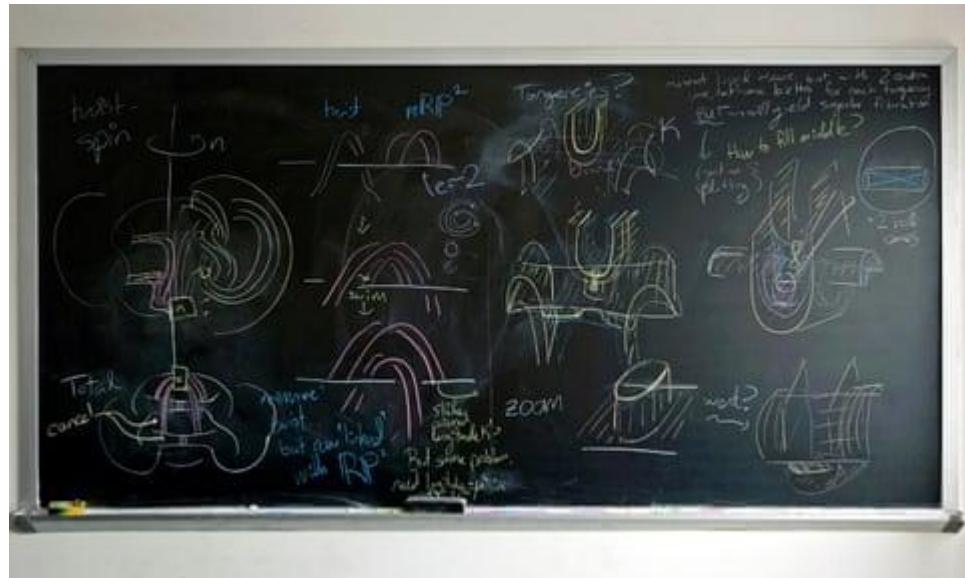


Figure 4. Spectra of the complex Lamé operator (3.1) for rectangular period lattices \mathcal{L} , $m = 1$ and $\omega = \omega_2$.

The Lamé equation

$$-\frac{d^2\psi}{dz^2} + m(m+1)\wp(z)\psi = \lambda\psi,$$

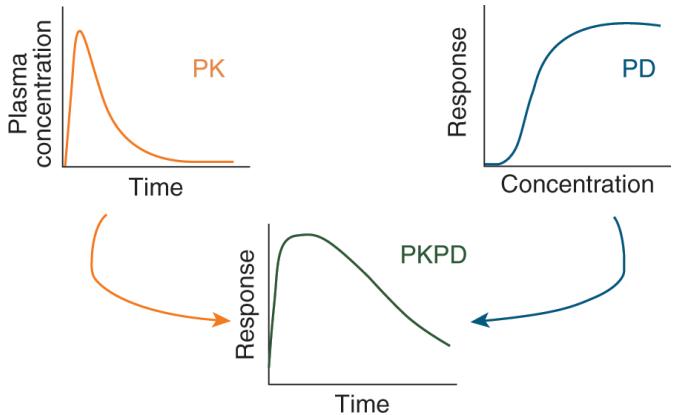


Mathematician Maggie Miller's blackboard at Princeton University. Photograph: Jessica Wynne
[Why mathematicians just can't quit their blackboards | Photography | The Guardian](#)

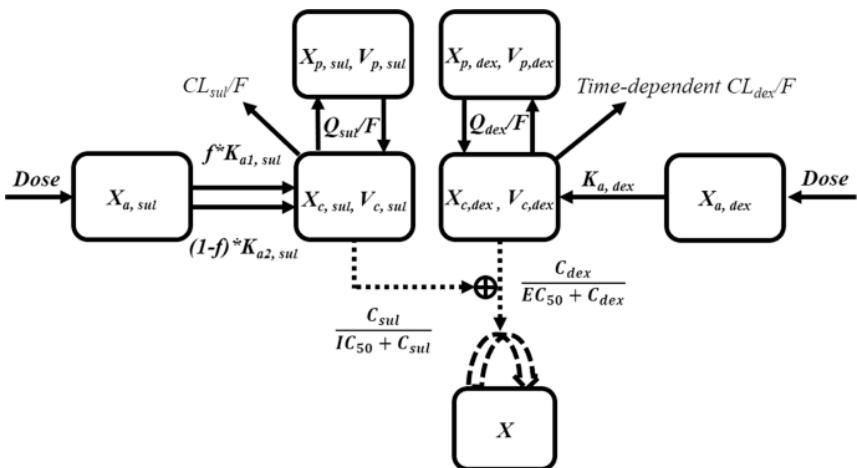


PK/PD modeller

2014-2017



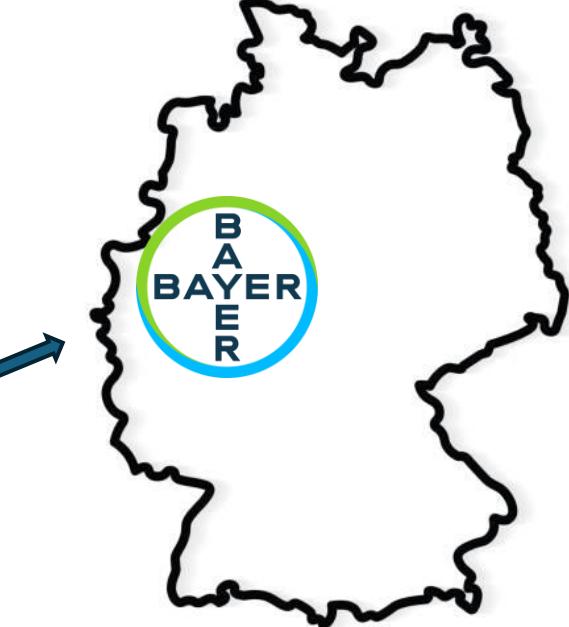
<https://aneskey.com/2-pharmacokinetic-and-pharmacodynamic-modelling-in-anaesthesia/>



<https://doi.org/10.1038/s41401-019-0251-7>



 **BAST**
Pharmaceutical
CRO



NONMEM®

The program for Nonlinear
Mixed Effects Modeling

```
$PROBLEM WARFARIN ; Problem title
$INPUT ID TIME AMT RATE EVID MDV DV ;Column names
$DATA warfarin.csv ignore=C

$SUBROUTINES ADVAN2 TRANS2 ; ADVAN2 is 1st order absorption, one central compartment
; MU referencing is not required, but can be helpful in evaluating FIM analytically,
; providing greater significant digit precision and speed
$PK
MU 1=LOG(THETA(1))
MU 2=LOG(THETA(2))
MU 3=LOG(THETA(3))
CL=EXP(MU_1+ETA(1))
V=EXP(MU_2+ETA(2))
KA=EXP(MU_3+ETA(3))
S2=V
F1=1.0

$ERROR
IPRED=A(2)/V
Y=IPRED + IPRED*EPS(1)

$THETA
0.15 ;[CL]
8.0 ;[V]
1.0 ;[KA]

$OMEGA (0.07) (0.02) (0.6)
$SIGMA 0.01 (0.001 FIXED)

;GROUPSIZE is a multiplier to the FIM
;FIMDIAG=1 means to use block diagonal FIM
$DESIGN GROUPSIZE=32 FIMDIAG=1
```

<https://doi.org/10.1002/psp.4.12713>

```
c:\Windows\system32\cmd.exe
ITERATION NO.: 5 OBJECTIVE VALUE: 258.957563934051 NO. OF FUNC.
EVALS.: 9
CUMULATIVE NO. OF FUNC. EVALS.: 50
PARAMETER: 7.0529E-01 2.2898E-01 3.1574E-01 1.3975E+00 1.1897E+00
GRADIENT: 9.589E+01 -8.2429E+00 3.1222E+00 3.8762E+01 3.2519E+01

ITERATION NO.: 10 OBJECTIVE VALUE: 239.964297920093 NO. OF FUNC.
EVALS.: 9
CUMULATIVE NO. OF FUNC. EVALS.: 92
PARAMETER: 8.3274E+00 1.0524E+01 1.1247E+01 2.7445E+01 4.3665E-02
GRADIENT: 6.1011E-01 1.5195E-01 2.1483E-01 1.1996E+00 8.3699E-01
S2=V
F1=1.0

$ERROR
IPRED=A(2)/V
Y=IPRED + IPRED*EPS(1)

$THETA
0.15 ;[CL]
8.0 ;[V]
1.0 ;[KA]

$OMEGA (0.07) (0.02) (0.6)
$SIGMA 0.01 (0.001 FIXED)

;GROUPSIZE is a multiplier to the FIM
;FIMDIAG=1 means to use block diagonal FIM
$DESIGN GROUPSIZE=32 FIMDIAG=1

Elapsed estimation time in seconds: 0.13
Elapsed covariance time in seconds: 0.05
Press any key to continue . . .
```

<https://www.pmxsolutions.com/2018/07/01/running-nonmem-models-and-r-scripts-from-batch-files/>

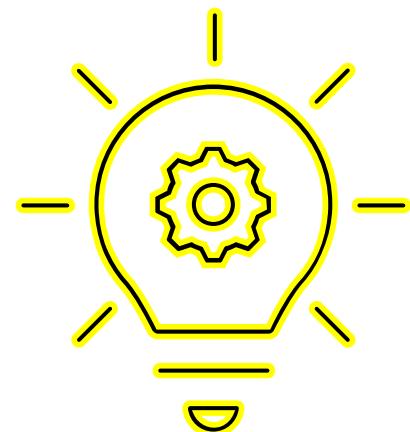
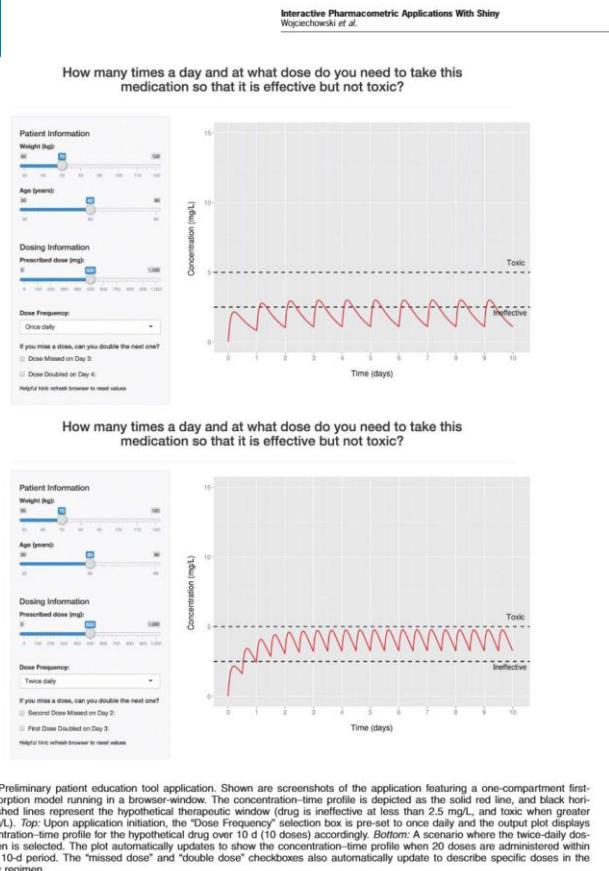
Intro to scripting/coding



```
library(nycflights13) ## package containing flights dataset
library(lubridate)
library(dplyr)
library(ggplot2)

head(flights, n = 3)
daily <- flights %>%
  mutate(date = make_date(year, month, day)) %>%
  count(date) %>%
  mutate(wday = wday(date, label = TRUE))
head(daily, n = 3)
ggplot(daily, aes(wday, n)) +
  geom_boxplot(outlier.colour = "hotpink") +
  labs(x = "Weekday", y = "Flights",
       subtitle = "Number of 2013 New York Flights Each Weekday")
#> # A tibble: 3 x 3
#>   year month day   n
#>   <dbl> <dbl> <dbl> <dbl>
#> 1 2013     1     1  517  515
#> 2 2013     1     2  533  529
#> 3 2013     1     3  542  540
#> # ... with 3 more variables: flight <int>, tailnum <chr>, origin <chr>, dest <chr>, air_time <dbl>,
#> # distance <dbl>, hour <dbl>, minute <dbl>, time_hour <dbl>,
#> > daily <- flights %>%
#> +   mutate(date = make_date(year, month, day)) %>%
#> +   count(date) %>%
#> +   mutate(wday = wday(date, label = TRUE))
#> > head(daily, n = 3)
#> # A tibble: 3 x 3
#>   date       n wday
#>   <date> <dbl> <dbl>
#> 1 2013-01-01 842 Tue
#> 2 2013-01-02 943 Wed
#> 3 2013-01-03 914 Thu
#> > ggplot(daily, aes(wday, n)) +
#> +   geom_boxplot(outlier.colour = "hotpink") +
#> +   labs(x = "Weekday", y = "Flights",
#> +         subtitle = "Number of 2013 New York Flights Each Weekday")
#>
```

cdhowe, CC BY-SA 4.0 <<https://creativecommons.org/licenses/by-sa/4.0/>>, via Wikimedia Commons



Personal programming projects



A screenshot of the PyCharm IDE interface. The main window shows a Python file named `tests.py` containing test cases for a "Polls" application. The tests include assertions for future questions, past questions, and two past questions. The code uses the `assertEqual` method to check if querysets are equal. The PyCharm interface includes toolbars, a sidebar with project files like `djip_first_steps`, `polls`, and `tests.py`, and a status bar at the bottom.

```
1 #!/usr/bin/env python
2 # coding: utf-8
3
4 from django.test import TestCase
5 from .models import Question
6
7 class QuestionViewTests(TestCase):
8     def test_index_view_with_no_questions(self):
9         """
10             If no questions exist, an appropriate message should be displayed.
11         """
12         response = self.client.get(reverse('polls:index'))
13         self.assertEqual(response.status_code, 200)
14         self.assertContains(response, "No polls are available")
15         self.assertQuerysetEqual(response.context['latest_question_list'], [])
16
17     def test_index_view_with_a_future_question(self):
18         """
19             Questions with a pub_date in the future should not be displayed on
20             the index page.
21         """
22         time = timezone.now() + timedelta(days=30)
23         future_question = Question(pub_date=time)
24         response = self.client.get(reverse('polls:index'))
25         self.assertEqual(response.status_code, 200)
26         self.assertContains(response, "No polls are available")
27         self.assertQuerysetEqual(response.context['latest_question_list'], [])
28
29     def test_index_view_with_no_past_question(self):
30         """
31             Questions with a pub_date in the past should be displayed on
32             the index page.
33         """
34         past_question = Question(pub_date=timezone.now() - timedelta(days=30))
35         response = self.client.get(reverse('polls:index'))
36         self.assertEqual(response.status_code, 200)
37         self.assertQuerysetEqual(response.context['latest_question_list'], [past_question])
38
39     def test_index_view_with_a_past_question(self):
40         """
41             Questions with a pub_date in the past should be displayed on
42             the index page.
43         """
44         time = timezone.now() - timedelta(days=30)
45         past_question = Question(pub_date=time)
46         response = self.client.get(reverse('polls:index'))
47         self.assertEqual(response.status_code, 200)
48         self.assertQuerysetEqual(response.context['latest_question_list'], [past_question])
49
50     def test_index_view_with_a_past_and_future_question(self):
51         """
52             Even if both past and future questions exist, only past questions
53             should be displayed.
54         """
55         past_question = Question(pub_date=timezone.now() - timedelta(days=30))
56         future_question = Question(pub_date=timezone.now() + timedelta(days=30))
57         response = self.client.get(reverse('polls:index'))
58         self.assertEqual(response.status_code, 200)
59         self.assertQuerysetEqual(response.context['latest_question_list'], [past_question])
60
61     def test_index_view_with_two_past_questions(self):
62         """
63             If there are multiple past questions, they should be displayed
64             in chronological order.
65         """
66         time = timezone.now() - timedelta(days=30)
67         first_question = Question(pub_date=time)
68         second_question = Question(pub_date=time + timedelta(days=1))
69         response = self.client.get(reverse('polls:index'))
70         self.assertEqual(response.status_code, 200)
71         self.assertQuerysetEqual(response.context['latest_question_list'], [second_question, first_question])
```

Copyright © 2000-2025 JetBrains s.r.o.

A screenshot of the [iMovies](#) website. The header features the text "i Movies" and "How many movies have you seen?". Below the header is a decorative film strip graphic with icons for "Movies", "Lists", and "People". The main content area is titled "All lists" and displays two items: "IMDb's Top 250" and "Reddit Top 250". Each item has a small thumbnail, a title, a description, and some statistics.

→ “MovieListMaker”

A screenshot of the [cohabio](#) website. The top navigation bar includes "Search again", "Download results report", and "Your Feedback". The main feature is a map of the UK showing travel times between Loughborough, UK and Northampton, UK. A callout box over the map provides specific travel time data: "Travel time for you: 48 mins", "Travel time for them: 41 mins (28 mins)", "Mean travel time: 44.5 mins", and "Standard deviation: 3.5 mins". Below the map are input fields for "Where you work:" (Loughborough, UK) and "Where they work:" (Northampton, UK). There are also sliders for "How long do you want to travel?" (90 mins) and "How long do they want to travel?" (60 mins), and dropdown menus for "Your preferred mode of transport" (Walk, Bike, Car, Public Transport) and "Their preferred mode of transport" (Walk, Bike, Car, Public Transport).

<https://cohabio.eu.pythonanywhere.com/>

<https://github.com/haessar/cohabio>

Search

Python dev/analyst in industry

2017 - 2019



Sector	Company	Location	Role	Duties
Digital marketing	Mavens of London (later Kantar)	Holborn, London	Dev team – Python dev	Building scraping and reporting tools
Sustainable energy	Energy Systems Catapult	Birmingham	Interim model analyst	Scripting for data processing and KPI calculation
Finance	JPMorgan Chase	Glasgow	Small cog in a corporate giant	ETL pipelines, Python 2 -> 3 migration



JPMC



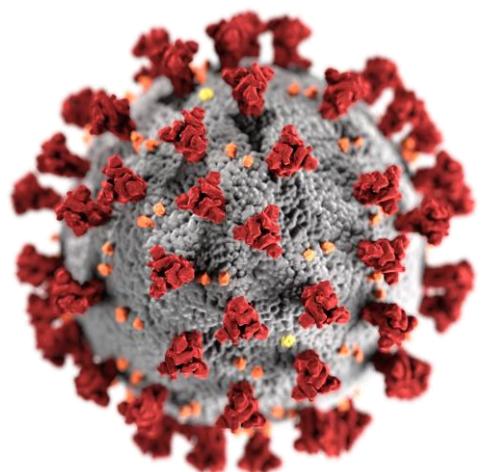
All code proprietary/private



3 months effectively unemployed

Covid and UofE

2020 - 2021



Alissa Eckert, MS; Dan Higgins, MAM, Public domain, via Wikimedia Commons



- Remote role
- Central IT rather than academic
- “People and Money” system - ETL pipelines
- Toxic team culture / worst job yet



- New to parenthood
- Isolation
- Low job satisfaction
- = Depression



Uni payment backlog left staff crying - report



THE UNIVERSITY
of EDINBURGH

Oliver Coe
BBC Scotland News

8 December 2023
Updated 9 December 2023

<https://www.bbc.co.uk/news/articles/cd1p32pd90po>

Entry to UofG

2021 - 2024



Prof Thomas Otto's lab



WCIP ISAB/Retreat 2022

COMPANION

Easy and reliable genome annotation.

Annotate your sequence!

- Open-source code
 - Publications
 - Collegiate culture

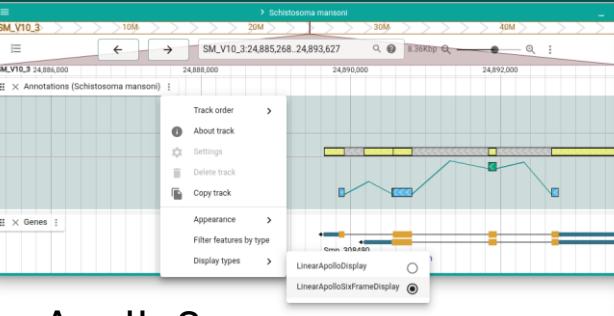
The figure displays the **peaks2utr** logo at the top, which includes a Python icon and a DNA helix icon. Below the logo is a plot titled "Forward-stranded 3' UTRs extended by several tools". The plot shows a chromatogram with vertical peaks representing gene expression levels across a genomic region. Colored bars above the chromatogram indicate different genomic features: light blue for gene, pink for reference 3' UTR, red for extended 3' UTR, purple for extended 3' UTR with SPAT, yellow for 3' UTR extended 3' UTR, and green for GETUTR extended 3' UTR. The x-axis at the bottom shows genomic coordinates from 499500 to 501500.

The figure shows the cellxgene Human Cell Landscape interface. On the left, a sidebar displays filters for batch, celltype, donor, louvain, and tissue. The tissue filter is expanded, showing categories like AdultAdipose, AdultAdrenalGland, and so on, each with a count (e.g., 1377, 23197, etc.). The main area features a 3D t-SNE plot where clusters of cells are represented as colored, semi-transparent shapes. To the right of the plot is a histogram titled 'Autosuggest genes' with two panels: one for 'n_genes' (ranging from 500 to 3k) and one for 'n_counts' (ranging from 100 to 10k). Below the plot, the Bioconductor logo is displayed with the text 'OPEN SOURCE SOFTWARE FOR BIOINFORMATICS'. In the bottom right corner, there are logos for paraCell, R, scanpy, and Seurat.

Current RSE role

2024 - Present

- Academic track – permanent contract
- Lots of autonomy
- Lots of variety
- = Job satisfaction



Apollo3

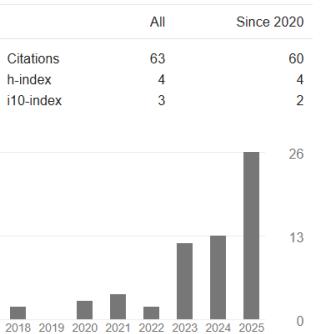
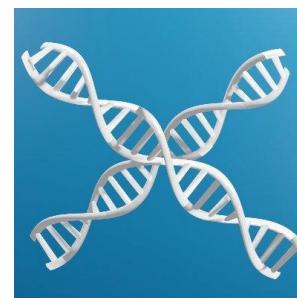
JBrowse

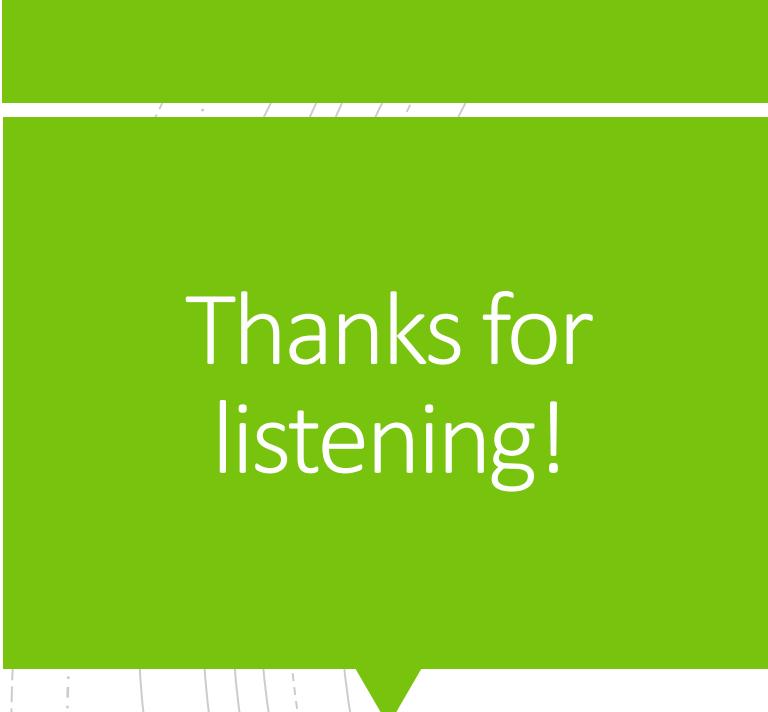
TS TypeScript

This block contains three logos: JBrowse (a stylized 'A' icon), Apollo3 (a screenshot of the software interface), and TypeScript (the 'TS' logo).



CompBio Hackathon 2025





Thanks for
listening!

“

My advice to others: don't be afraid to try a variety of roles before settling on something that works for you. There will always be opportunities arising in unexpected areas, and doors reopening that once might have seemed firmly shut.

”

<https://theauditorium.blog/2025/03/12/pathfinder-career-narratives-60-research-software-engineer/>