

P&S Accelerating Genomics

Introduction & Course Logistics

Dr. Mohammed Alser

 @mealsr

ETH Zurich

Fall 2022

27 October 2022

SAFARI

ETH zürich

Intelligent Genomic Analyses

Intelligent Genomic Analyses (Fall 2022)

<https://www.youtube.com/watch?v=nA41964-9r8>

What is Intelligent Genome Analysis?

- Fast genome analysis

- Real-time analysis?*

Bandwidth

- Population-scale genome analysis

- Number of analyses per day!*

Scalability

- Using intelligent architectures

- Small specialized HW with less data movement*

Energy-efficiency &
Portability

- DNA is a valuable asset

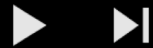
- Controlled-access analysis*

Privacy

- Avoiding erroneous analysis

- E.g., your father is not your father*

Accuracy



SAFARI

20:13 / 2:44:20



Intelligent Genome Analysis

Mohammed Alser, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu

[“From Molecules to Genomic Variations: Intelligent Algorithms and Architectures for Intelligent Genome Analysis”](#)

Computational and Structural Biotechnology Journal, 2022

[\[Source code\]](#)



ELSEVIER



journal homepage: www.elsevier.com/locate/csbj



Review

From molecules to genomic variations: Accelerating genome analysis via intelligent algorithms and architectures



Mohammed Alser*, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu*

ETH Zurich, Gloriastrasse 35, 8092 Zürich, Switzerland

Two P&S Genomics Courses

P&S Mobile Genomics Introduction & Project Proposals

Dr. Mohammed Alser
@mealser

ETH Zurich
PLAY ALL
8 March 2022


Livestream - P&S Genome Sequencing on Mobile Devices (Spring 2022)

11 videos • 213 views • Last updated on 5 Jun 2022




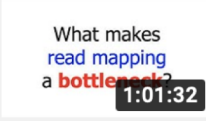
Onur Mutlu Lectures

SUBSCRIBE


- 

P&S Mobile Genomics
Introduction & Project Proposals
Dr. Mohammed Alser
ETH Zurich
Spring 2022
8 March 2022
28:26


Mobile Genomics Course - Meeting 1: Course Introduction & Project Proposals (Spring 2022)
Onur Mutlu Lectures
- 

Mobile Genomics Course - Meeting 2: Introduction to Sequencing (Spring 2022)
Onur Mutlu Lectures
- 

What makes
read mapping
a **bottle neck**?
1:01:32

Mobile Genomics Course - Meeting 3: Read Mapping (Spring 2022)
Onur Mutlu Lectures
- 

Process whole genome at 30x coverage in \approx25 minutes with hardware support for data compression
1:01:25

Mobile Genomics Course - Meeting 4: GateKeeper (Spring 2022)
Onur Mutlu Lectures
- 

P&S Mobile Genomics
Lecture 5: MAGNET & Shouji
Dr. Mohammed Alser
ETH Zurich
Spring 2022
4 April 2022
54:39

Mobile Genomics Course - Meeting 5: MAGNET & Shouji (Spring 2022)
Onur Mutlu Lectures

https://www.youtube.com/playlist?list=PL5Q2soXY2Zi8NrPDgOR1yRU_Cxxjw-u18

Genomics (Spring 2022)

Spring 2022 Edition:

- https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=genome_seq_mobile

Youtube Livestream:

- https://www.youtube.com/playlist?list=PL5Q2soXY2Zi8NrPDgOR1yRU_Cxxjw-u18

Project course

- Taken by Bachelor's/Master's students
- Genomics lectures
- Hands-on research exploration
- Many research readings

Mobile Genomics Course - Meeting 1: Course Introduction... Watch later 1/11

P&S Mobile Genomics

Introduction & Project Proposals

Dr. Mohammed Alser
@mealsar

ETH Zurich
Spring 2022
8 March 2022

Watch on

ETH zürich

Spring 2022 Meetings/Schedule

Week	Date	Livestream	Meeting	Learning Materials	Assignments
W1	8.3 Tue.	Live	M1: P&S Mobile Genomics Course Introduction & Project Proposals (PDF) (PPT)	Required Materials Recommended Materials	
W2	15.3 Tue.	Live	M2: Introduction to Sequencing (PDF) (PPT)		
W3	22.3 Tue.	Live	M3: Read Mapping (PDF) (PPT)		
W4	29.3 Tue.	Live	M4: GateKeeper (PDF) (PPT)		
W5	5.4 Tue.	Live	M5: MAGNET & Shouji (PDF) (PPT) Video		
W6	12.4 Tue.	Live	M6: SneakySnake (PDF) (PPT)		
W7	26.4 Tue.	Live	M7: GenStore (PDF) (PPT)		
W8	03.05 Tue.	Premiere	M8: GRIM-Filter (PDF) (PPT)		
W9	10.05 Tue.	Live	M9: Genome Assembly (PDF) (PPT)		
W10	17.05 Tue.	Live	M10: Genomic Data Sharing Under Differential Privacy (PDF) (PPT)		
W11	07.06 Tue.	Premiere	M11: Accelerating Genome Sequence Analysis (PDF) (PPT)		

Course Materials

Spring 2022 Meetings/Schedule

Week	Date	Livestream	Meeting	Learning Materials	Assignments
W1	8.3 Tue.	Live	M1: P&S Mobile Genomics Course Introduction & Project Proposals (PDF) (PPT)	Required Materials Recommended Materials	
W2	15.3 Tue.	Live	M2: Introduction to Sequencing (PDF) (PPT)		
W3	22.3 Tue.	Live	M3: Read Mapping (PDF) (PPT)		
W4	29.3 Tue.	Live	M4: GateKeeper (PDF) (PPT)		
W5	5.4 Tue.	Live	M5: MAGNET & Shouji (PDF) (PPT) Video		
W6	12.4 Tue.	Live	M6: SneakySnake (PDF) (PPT)		
W7	26.4 Tue.	Live	M7: GenStore (PDF) (PPT)		
W8	03.05 Tue.	Premiere	M8: GRIM-Filter (PDF) (PPT)		
W9	10.05 Tue.	Live	M9: Genome Assembly (PDF) (PPT)		
W10	17.05 Tue.	Live	M10: Genomic Data Sharing Under Differential Privacy (PDF) (PPT)		
W11	07.06 Tue.	Premiere	M11: Accelerating Genome Sequence Analysis (PDF) (PPT)		

https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=genome_seq_mobile

Past Semesters (P&S Accelerating Genomics)

- **Fall 2022:**

https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=bioinformatics

- **Spring 2022:**

https://safari.ethz.ch/projects_and_seminars/spring2022/doku.php?id=bioinformatics

- **Fall 2021:**

https://safari.ethz.ch/projects_and_seminars/fall2021/doku.php?id=bioinformatics

- **Spring 2021:**

https://safari.ethz.ch/projects_and_seminars/spring2021/doku.php?id=bioinformatics

- **Fall 2020:**

https://safari.ethz.ch/projects_and_seminars/fall2020/doku.php?id=bioinformatics

Past Semesters (P&S Mobile Genomics)

- **Fall 2022:**

https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=genome_seq_mobile

- **Spring 2022:**

https://safari.ethz.ch/projects_and_seminars/spring2022/doku.php?id=genome_seq_mobile

- **Fall 2021:**

https://safari.ethz.ch/projects_and_seminars/fall2021/doku.php?id=genome_seq_mobile

- **Spring 2021:**

https://safari.ethz.ch/projects_and_seminars/spring2021/doku.php?id=genome_seq_mobile

- **Fall 2020:**

https://safari.ethz.ch/projects_and_seminars/fall2020/doku.php?id=genome_seq_mobile

The Role of This Course

Projects & Seminars: Accelerating Genomics

- We will cover the **basics** of **genome analysis** to understand the **computational steps** of the entire pipeline and find the computational bottlenecks.
- Students will learn about the **existing efforts** for **accelerating** one or more of these steps.
- Student will carry out a **hands-on project** to implement and improve these acceleration efforts.

Key Objectives

- Multiple components that are aimed at improving students'
 - **Basic knowledge** in genome analysis (**dry lab**)
 - **Technical skills** in genome analysis and computer architecture
 - **Critical thinking and analysis**
 - **Familiarity with key research directions**
 - **Technical presentation** of your project

Key Goal

(Learn how to)

efficiently implement

one of the key steps in genome

analysis on portable devices


Prerequisites of the Course

- No prior knowledge in bioinformatics or genome analysis is required.
- Digital Design and Computer Architecture (or equivalent course)
- A good knowledge in C programming language is required.
- Experience in at least one of the following is highly desirable: FPGA implementation and GPU programming.
- Interest in making things efficient and solving problems

Course Info: Who Are We?



Mohammed Alser

- Lecturer and Senior Researcher, [SAFARI Research Group](#), [ETH Zürich](#), since Sept. 2018.
- PhD from Bilkent University (Turkey) 2018, worked at UCLA, TU Dresden, and PETRONAS.
- [Received the IEEE Turkey Doctoral Dissertation Award](#) and a number of international prestigious awards.
-  <https://twitter.com/mealser>
- My main research is in **bioinformatics, computational genomics, metagenomics**, and computer architecture.
- I am especially excited about **building** new data structures, algorithms, and architectures that **make intelligent genome analysis a reality**.

Course Info: Who Are We? (I)



Juan Gómez Luna

Senior Researcher and
Lecturer

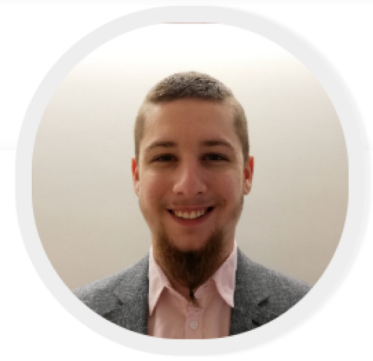
Processing-In-Memory |
Heterogeneous computing |
Memory Systems | Bioinformatics |
Medical imaging



Can Firtina

PhD Student

Genome Assembly |
Sequence Analysis &
Alignment | Biologically-
Inspired Computing
Paradigms | Brain-
Computer Interfaces |
Phase-change memory



Joël Lindegger

PhD Student

Acceleration of the
bioinformatics pipeline |
Current and future
computer architectures |
All kinds of algorithms
and data structures

Course Info: Who Are We? (II)



**Nika
Mansourighiasi**

PhD Student

Processing-In-Memory |
Emerging Memory &
Processing Technologies



Max Rumpf
Masters Student

Bioinformatics | Computational
Genomics | Sequence Analysis &
Alignment | Machine Learning



Julien Eudine
Masters Student

Computational Genomics
| High-Performance
Computing |
Heterogeneous
Computing Systems |
Hardware acceleration

Course Info: Who Are We? (III)



Arvid Gollwitzer

Masters Student

Bioinformatics |
Computational Genomics
| Sequence Analysis &
Alignment | Medical
Applications | Clinical
Metagenomics



Younjoo Lee

Masters Student

Computer architecture |
FPGA-acceleration |
Bioinformatics |
Processing in Memory.



Luca Blum

Masters Student

Machine Learning |
Computational Genomics
| Bioinformatics | High-
Performance Computing.

Course Requirements and Expectations

- Attendance required for all meetings
- Study the learning materials
- Each student will carry out a hands-on project
 - Build, implement, code, and design with close engagement from the supervisors
- Participation
 - Ask questions, contribute thoughts/ideas
 - Read relevant papers
- Presentation & GitHub repository

We will help the projects with good progress to get published in good venues!

Course Website

- https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=bioinformatics
- Useful information for the course
- We will also use Glip, Moodle, and E-mail for Q&A, announcements, ..
- Check your Glip, Moodle, and E-mail frequently for announcements

Project Assignment

- We ask you to **study the learning materials** before next meetings!
- We gave you **5 days** to enter your preferences,
- Then, we **match** your interests, skills, and background with a suitable project.
- We assign you a project and put you in touch with a **mentor** in the second week.

Assignment 1: Intelligent Genome Analysis

Mohammed Alser, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu

[“From Molecules to Genomic Variations: Intelligent Algorithms and Architectures for Intelligent Genome Analysis”](#)

Computational and Structural Biotechnology Journal, 2022

[\[Source code\]](#)



ELSEVIER



journal homepage: www.elsevier.com/locate/csbj



Review

From molecules to genomic variations: Accelerating genome analysis via intelligent algorithms and architectures



Mohammed Alser*, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu*

ETH Zurich, Gloriastrasse 35, 8092 Zürich, Switzerland

Assignment 2: Basic Information

Basic Information

P&S Genomics - Fall 2022 @ ETH Zurich

We would like to warmly welcome all of you to the *P&S Genomics course* :)

Because we care about you, we would like to know more about you and your preferences to serve you better in this course and plan for a fruitful course moving forward ;)

The questionnaire is due on 4 October 2022, 23:59 Zurich time.

 mealser@gmail.com (not shared) [Switch account](#)

 Draft saved

* Required

Full name *

Your answer

 This is a required question

Email address *

Your answer

Assignment 2: Project Preferences

Which of the following technical skills do you have? *

Choose what you feel comfortable to use throughout your project. Choose all applicable.

	Stay away	Basic knowledge (only took a course)	Good knowledge (did a few projects)	Did several projects	Expert
C/Verilog/VHDL programming for FPGAs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
CUDA programming for GPUs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C programming	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C++ programming	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Python programming	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Surveying a topic & benchmarking tools	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am fine with using any programming language and learning new skills	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Your Responsibilities

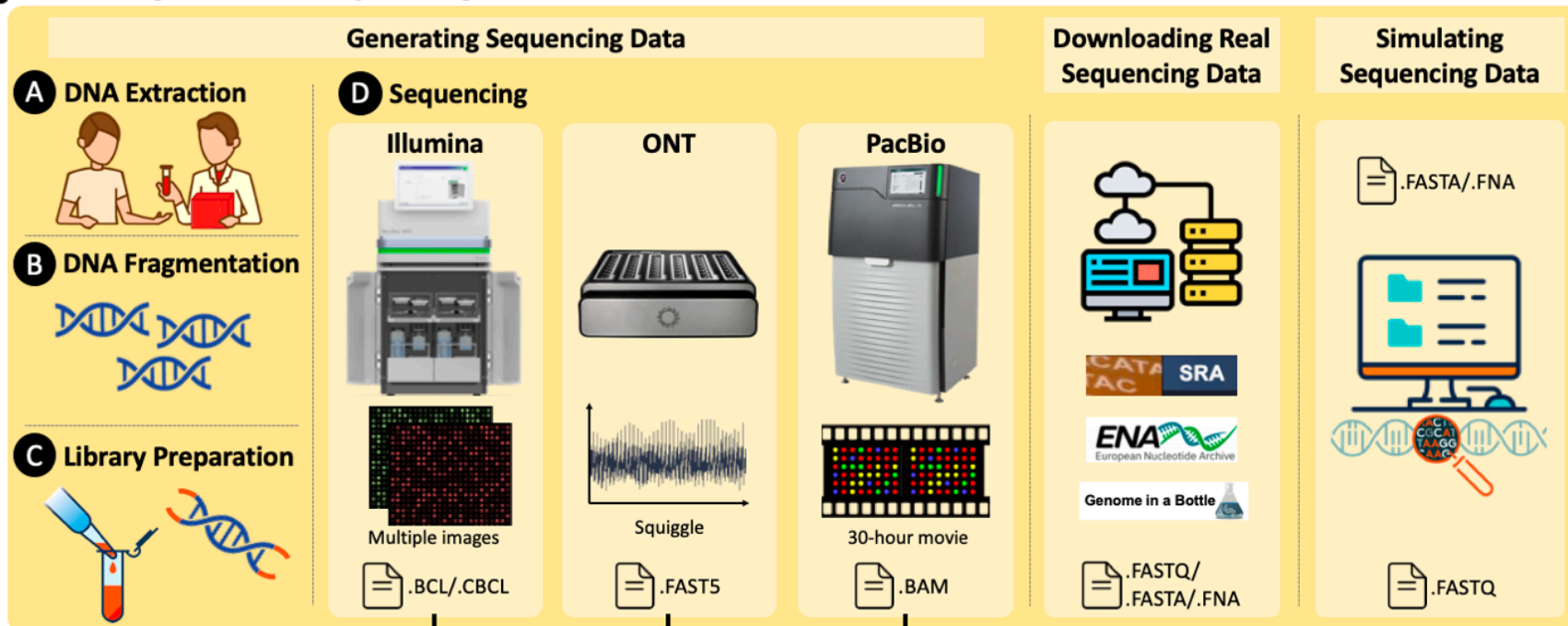
- 1 Lecture every week
 - Monday 4-5 PM
 - https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=genome_seq_mobile
 - Thursday 10-11 AM
 - https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=bioinformatics
- Attendance is mandatory
- Working on your project for ~6 hours per week
- Meeting your mentors weekly is required
- Replying on time to Glip, Moodle, and E-mail messages.

What is Intelligent Genome Analysis?

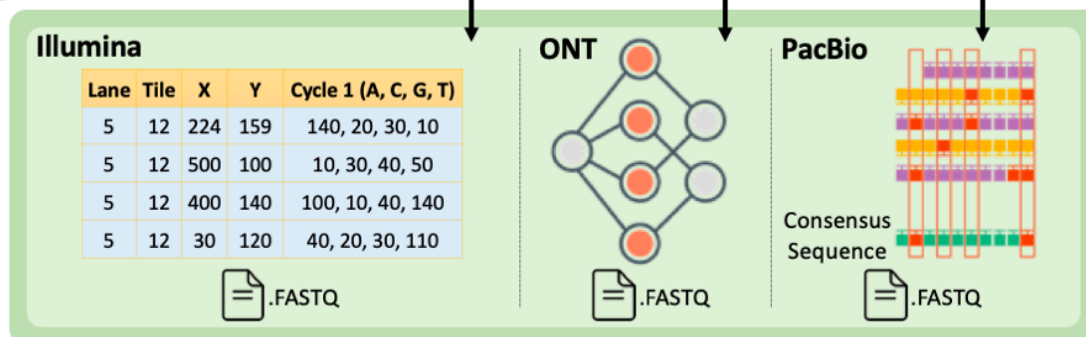
- Fast genome analysis Bandwidth
 - *Real-time analysis*
- Using intelligent architectures Energy-efficiency & Latency
 - *Specialized HW with less data movement*
- DNA is a valuable asset Privacy
 - *Controlled-access analysis*
- Population-scale genome analysis Scalability
 - *Sequence anywhere at large scale!*
- Avoiding erroneous analysis Accuracy
 - *E.g., your father is not your father*

Next Meetings

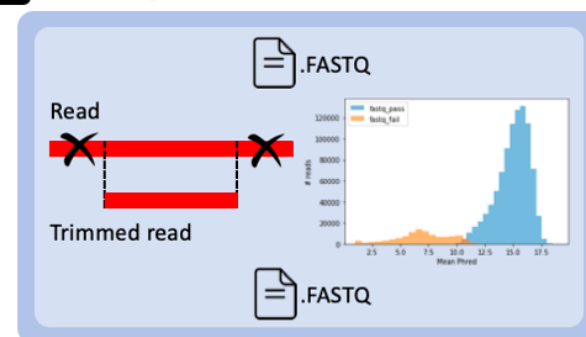
1 Obtaining Genomic Sequencing Data



2 Basecalling

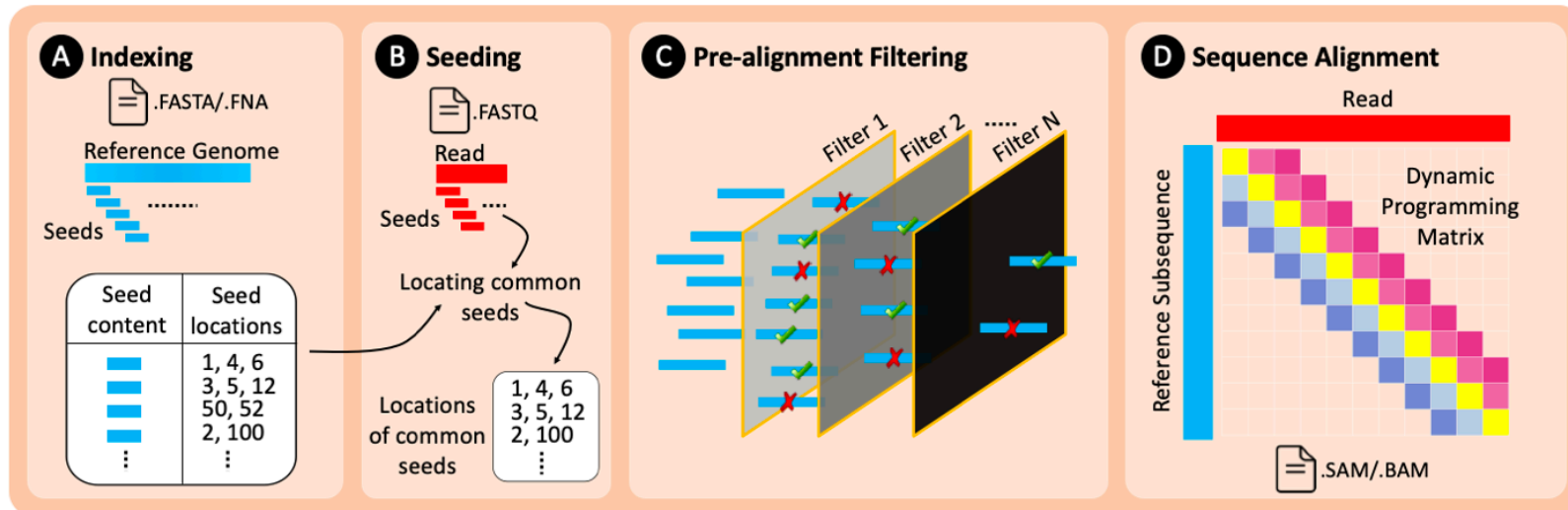


3 Quality Control

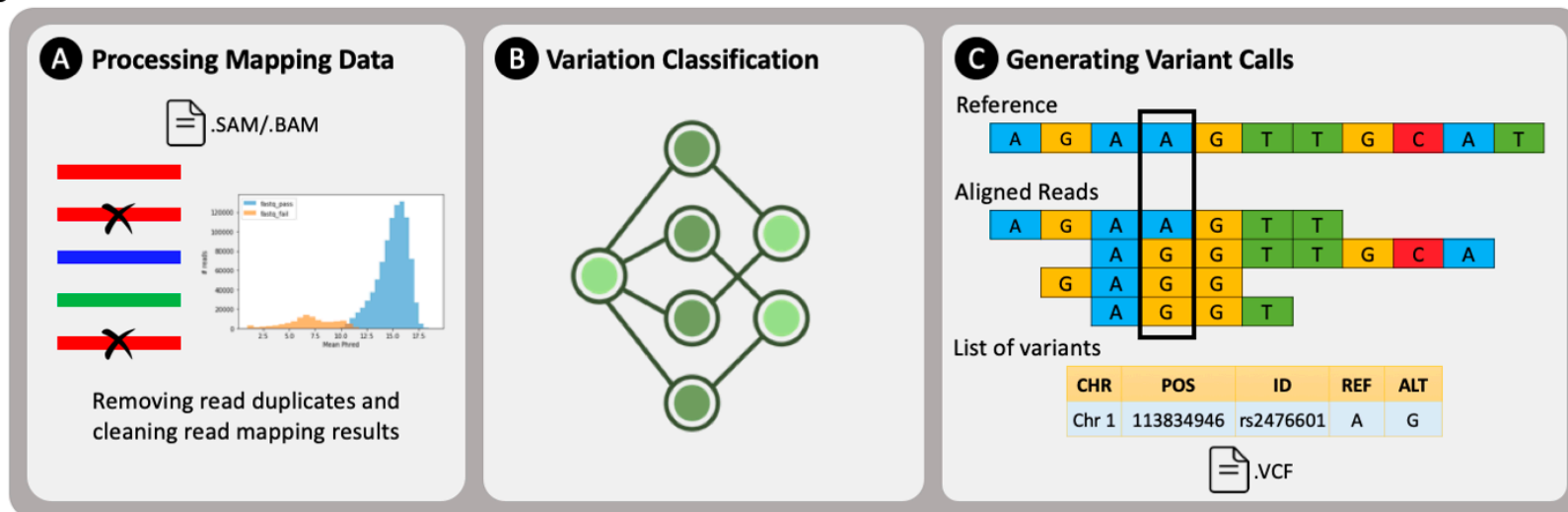


Next Meetings

4 Read Mapping



5 Variant Calling



Next Meetings

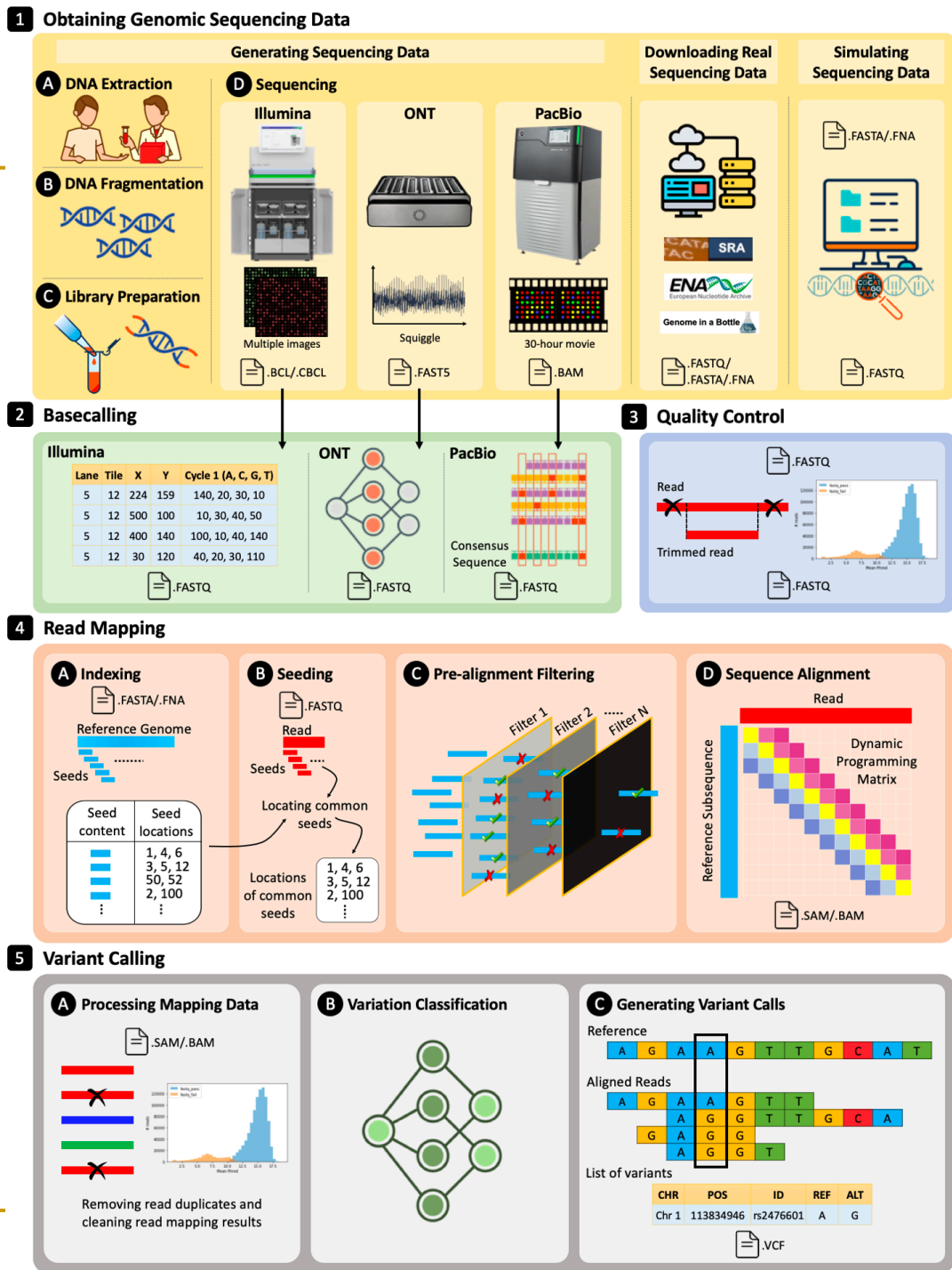
- 1 Lecture every week

Monday 4-5 PM

- https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=genome_seq_mobile

Thursday 10-11 AM

- https://safari.ethz.ch/projects_and_seminars/fall2022/doku.php?id=bioinformatics



A Indexing

.FASTA/.FNA

Reference Genome

Seeds

Seed content	Seed locations
1, 4, 6	1, 4, 6
3, 5, 12	3, 5, 12
50, 52	50, 52
2, 100	2, 100
⋮	⋮

B Seeding

.FASTQ

Read

Seeds

Locating common seeds

Locations of common seeds

1, 4, 6
3, 5, 12
2, 100
⋮

C Pre-alignment Filtering

D Sequence Alignment

Read

Reference Subsequence

Dynamic Programming Matrix

.SAM/.BAM

A Processing Mapping Data

.SAM/.BAM

Removing read duplicates and cleaning read mapping results

B Variation Classification

C Generating Variant Calls

Reference

A G A A G T T G C A T

Aligned Reads

A G A A G T T

A A G G T T G C A

G A G G

A G G T

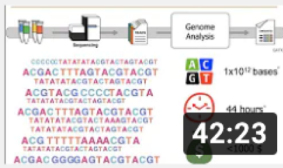
List of variants

CHR	POS	ID	REF	ALT
Chr 1	113834946	rs2476601	A	G

.VCF

Topics To Be Covered (I)

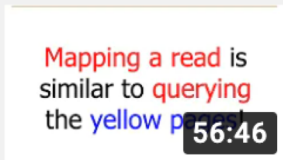
3



Mobile Genomics Course - Meeting 2: Introduction to Sequencing (Fall 2021)

Onur Mutlu Lectures

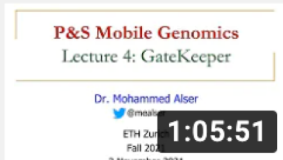
4



Mobile Genomics Course - Meeting 3: Read Mapping (Fall 2021)

Onur Mutlu Lectures

5



Mobile Genomics Course - Meeting 4: GateKeeper (Fall 2021)

Onur Mutlu Lectures

6



Mobile Genomics Course - Lecture 5: MAGNET & Shouji (Fall 2021)

Onur Mutlu Lectures

7



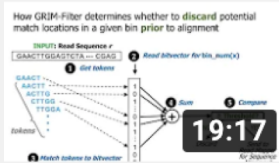
Computer Architecture - Lecture 8: Intelligent Genome Analysis (ETH Zürich, Fall 2020)

Onur Mutlu Lectures

⋮

Topics To Be Covered (II)

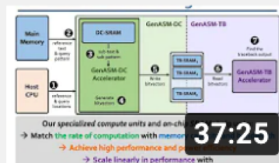
8



GRIM-Filter: Fast Seed Location Filtering in DNA Read Mapping w/ Processing-in-Memory - Jeremie Kim

Onur Mutlu Lectures

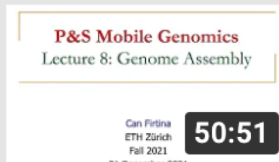
9



Comp. Architecture - Lecture 9a: GenASM: Approx. String Matching Accelerator (ETH Zürich, Fall 2020)

Onur Mutlu Lectures

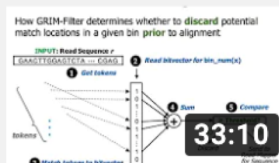
10



Mobile Genomics Course - Lecture 8: Genome Assembly (Fall 2021)

Onur Mutlu Lectures

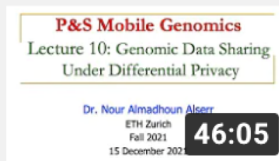
11



Mobile Genomics Course - Lecture 9: GRIM-Filter (Fall 2021)

Onur Mutlu Lectures

12



Mobile Genomics Course - Lecture 10: Genomic Data Sharing Under Differential Privacy (Fall 2021)

Onur Mutlu Lectures

⋮

Intelligent Genome Analysis

Mohammed Alser, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu

[“From Molecules to Genomic Variations: Intelligent Algorithms and Architectures for Intelligent Genome Analysis”](#)

Computational and Structural Biotechnology Journal, 2022

[\[Source code\]](#)



ELSEVIER



journal homepage: www.elsevier.com/locate/csbj



Review

From molecules to genomic variations: Accelerating genome analysis via intelligent algorithms and architectures



Mohammed Alser*, Joel Lindegger, Can Firtina, Nour Almadhoun, Haiyu Mao, Gagandeep Singh, Juan Gomez-Luna, Onur Mutlu*

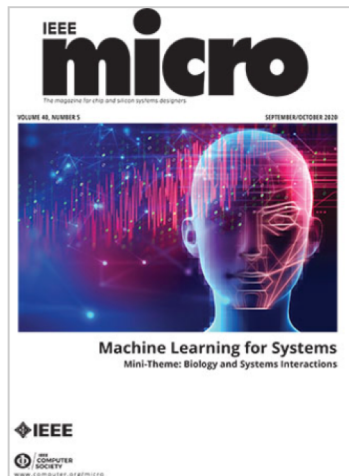
ETH Zurich, Gloriastrasse 35, 8092 Zürich, Switzerland

Near-memory Pre-alignment Filtering

Mohammed Alser, Zülal Bingöl, Damla Senol Cali, Jeremie Kim, Saugata Ghose, Can Alkan, Onur Mutlu

[“Accelerating Genome Analysis: A Primer on an Ongoing Journey”](#)

IEEE Micro, August 2020.



[Home](#) / [Magazines](#) / [IEEE Micro](#) / 2020.05

IEEE Micro

Accelerating Genome Analysis: A Primer on an Ongoing Journey

Sept.-Oct. 2020, pp. 65-75, vol. 40

DOI Bookmark: [10.1109/MM.2020.3013728](https://doi.org/10.1109/MM.2020.3013728)

Authors

[Mohammed Alser](#), ETH Zürich

[Zulal Bingol](#), Bilkent University

[Damla Senol Cali](#), Carnegie Mellon University

[Jeremie Kim](#), ETH Zurich and Carnegie Mellon University

[Saugata Ghose](#), University of Illinois at Urbana-Champaign and Carnegie Mellon University

[Can Alkan](#), Bilkent University

[Onur Mutlu](#), ETH Zurich, Carnegie Mellon University, and Bilkent University

◀	▶
Previous	Next
☰	Table of Contents
📄	Past Issues

Read Mapping in 111 pages!

In-depth analysis of 107 read mappers (1988-2020)

Mohammed Alser, Jeremy Rotman, Dhrithi Deshpande, Kodi Taraszka, Huwenbo Shi, Pelin Icer Baykal, Harry Taegyung Yang, Victor Xue, Sergey Knyazev, Benjamin D. Singer, Brunilda Balliu, David Koslicki, Pavel Skums, Alex Zelikovsky, Can Alkan, Onur Mutlu, Serghei Mangul

["Technology dictates algorithms: Recent developments in read alignment"](#)

Genome Biology, 2021

[\[Source code\]](#)

Alser et al. *Genome Biology* (2021) 22:249
<https://doi.org/10.1186/s13059-021-02443-7>


Genome Biology

REVIEW

Open Access

Technology dictates algorithms: recent developments in read alignment



Mohammed Alser^{1,2,3†}, Jeremy Rotman^{4†}, Dhrithi Deshpande⁵, Kodi Taraszka⁴, Huwenbo Shi^{6,7}, Pelin Icer Baykal⁸, Harry Taegyung Yang^{4,9}, Victor Xue⁴, Sergey Knyazev⁸, Benjamin D. Singer^{10,11,12}, Brunilda Balliu¹³, David Koslicki^{14,15,16}, Pavel Skums⁸, Alex Zelikovsky^{8,17}, Can Alkan^{2,18}, Onur Mutlu^{1,2,3†} and Serghei Mangul^{5*†} 

Feedback From Our Community!



James Ferguson

@Psy_Fer_

This is awesome! I've got my evening reading sorted.



Stéphane Le Crom

@slecrom

Very complete article on the evolution of read alignment algorithms. [#NGS](#) [#genomics](#)



Svetlana Gorokhova

@SGorokhova

An impressive overview of read alignment methods over the last three decades



BContrerasMoreira @BrunoContrerasM · Sep 10

Replying to [@meals](#) [@GenomeBiology](#) and 3 others

Buen hilo de repaso sobre la evolución de los algoritmos de alineamiento de secuencias a medida que ha mejorado la tecnología de secuenciación

<https://twitter.com/meals/status/1435223377644503040>

GenASM Framework [MICRO 2020]

Damla Senol Cali, Gurpreet S. Kalsi, Zülal Bingöl, Can Firtina, Lavanya Subramanian, Jeremie S. Kim, Rachata Ausavarungnirun, Mohammed Alser, Juan Gomez-Luna, Amirali Boroumand, Anant Nori, Allison Scibisz, Sreenivas Subramoney, Can Alkan, Saugata Ghose, and Onur Mutlu,

"GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis"

Proceedings of the [53rd International Symposium on Microarchitecture \(MICRO\)](#), Virtual, October 2020.

[[Lightning Talk Video](#) (1.5 minutes)]

[[Lightning Talk Slides \(pptx\)](#) ([pdf](#))]

[[Talk Video](#) (18 minutes)]

[[Slides \(pptx\)](#) ([pdf](#))]

GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis

Damla Senol Cali[†]✕ Gurpreet S. Kalsi[✕] Zülal Bingöl[∇] Can Firtina[◇] Lavanya Subramanian[‡] Jeremie S. Kim[◇]†
Rachata Ausavarungnirun[○] Mohammed Alser[◇] Juan Gomez-Luna[◇] Amirali Boroumand[†] Anant Nori[✕]
Allison Scibisz[†] Sreenivas Subramoney[✕] Can Alkan[∇] Saugata Ghose^{*†} Onur Mutlu[◇]†∇

[†]Carnegie Mellon University [✕]Processor Architecture Research Lab, Intel Labs [∇]Bilkent University [◇]ETH Zürich

[‡]Facebook [○]King Mongkut's University of Technology North Bangkok ^{*}University of Illinois at Urbana-Champaign

GenStore (ASPLOS 2022)

Nika Mansouri Ghiasi, Jisung Park, Harun Mustafa, Jeremie Kim, Ataberk Olgun, Arvid Gollwitzer, Damla Senol Cali, Can Firtina, Haiyu Mao, Nour Almadhoun Alserr, Rachata Ausavarungnirun, Nandita Vijaykumar, **Mohammed Alser**, Onur Mutlu
["GenStore: A High-Performance and Energy-Efficient In-Storage Computing System for Genome Sequence Analysis"](#),
ASPLOS 2022

GenStore: A High-Performance In-Storage Processing System for Genome Sequence Analysis

Nika Mansouri Ghiasi
ETH Zürich
Switzerland

Jisung Park
ETH Zürich
Switzerland

Harun Mustafa
ETH Zürich
Switzerland

Jeremie Kim
ETH Zürich
Switzerland

Ataberk Olgun
ETH Zürich
Switzerland

Arvid Gollwitzer
ETH Zürich
Switzerland

Damla Senol Cali
Bionano Genomics
USA

Can Firtina
ETH Zürich
Switzerland

Haiyu Mao
ETH Zürich
Switzerland

Nour Almadhoun
Alserr
ETH Zürich
Switzerland

Rachata
Ausavarungnirun
KMUTNB
Thailand

Nandita Vijaykumar
University of Toronto
Canada

Mohammed Alser
ETH Zürich
Switzerland

Onur Mutlu
ETH Zürich
Switzerland

GenPIP (MICRO 2022)

Haiyu Mao, **Mohammed Alser**, Mohammad Sadrosadati, Can Firtina, Akanksha Baranwal, Damla Senol Cali, Aditya Manglik, Nour Almadhoun Alserr, Onur Mutlu

[“GenPIP: In-Memory Acceleration of Genome Analysis via Tight Integration of Basecalling and Read Mapping”](#)

Proceedings of the [55rd International Symposium on Microarchitecture \(MICRO\)](#), 2022.

GenPIP: In-Memory Acceleration of Genome Analysis via Tight Integration of Basecalling and Read Mapping

Haiyu Mao¹ Mohammed Alser¹ Mohammad Sadrosadati¹ Can Firtina¹ Akanksha Baranwal¹
Damla Senol Cali² Aditya Manglik¹ Nour Almadhoun Alserr¹ Onur Mutlu¹

¹*ETH Zürich*

²*Bionano Genomics*

SeGraM (ISCA 2022)

Damla Senol Cali, Konstantinos Kanellopoulos, Joel Lindegger, Zülal Bingöl, Gurpreet S. Kalsi, Ziyi Zuo, Can Firtina, Meryem Banu Cavlak, Jeremie Kim, Nika Mansouri Ghiasi, Gagandeep Singh, Juan Gómez-Luna, Nour Almadhoun Alserr,

Mohammed Alser, Sreenivas Subramoney, Can Alkan, Saugata Ghose, Onur Mutlu
[“SeGraM: A Universal Hardware Accelerator for Genomic Sequence-to-Graph and Sequence-to-Sequence Mapping”](#)

ISCA 2022

SeGraM: A Universal Hardware Accelerator for Genomic Sequence-to-Graph and Sequence-to-Sequence Mapping

Damla Senol Cali¹ Konstantinos Kanellopoulos² Joël Lindegger² Zülal Bingöl³
Gurpreet S. Kalsi⁴ Ziyi Zuo⁵ Can Firtina² Meryem Banu Cavlak² Jeremie Kim²
Nika Mansouri Ghiasi² Gagandeep Singh² Juan Gómez-Luna² Nour Almadhoun Alserr²
Mohammed Alser² Sreenivas Subramoney⁴ Can Alkan³ Saugata Ghose⁶ Onur Mutlu²

¹Bionano Genomics ²ETH Zürich ³Bilkent University ⁴Intel Labs
⁵Carnegie Mellon University ⁶University of Illinois Urbana-Champaign

GenASM Framework [MICRO 2020]

- Damla Senol Cali, Gurpreet S. Kalsi, Zülal Bingöl, Can Firtina, Lavanya Subramanian, Jeremie S. Kim, Rachata Ausavarungnirun, **Mohammed Alser**, Juan Gomez-Luna, Amirali Boroumand, Anant Nori, Allison Scibisz, Sreenivas Subramoney, Can Alkan, Saugata Ghose, and Onur Mutlu, "[GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis](#)"
Proceedings of the [53rd International Symposium on Microarchitecture \(MICRO\)](#), Virtual, October 2020.
[[Lightning Talk Video](#) (1.5 minutes)]
[[Lightning Talk Slides \(pptx\)](#) ([pdf](#))]
[[Talk Video](#) (18 minutes)]
[[Slides \(pptx\)](#) ([pdf](#))]

GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis

Damla Senol Cali^{†✕} Gurpreet S. Kalsi[✕] Zülal Bingöl[∇] Can Firtina[◇] Lavanya Subramanian[‡] Jeremie S. Kim^{◇†}
Rachata Ausavarungnirun[○] Mohammed Alser[◇] Juan Gomez-Luna[◇] Amirali Boroumand[†] Anant Nori[✕]
Allison Scibisz[†] Sreenivas Subramoney[✕] Can Alkan[∇] Saugata Ghose^{*†} Onur Mutlu^{◇†∇}
[†]Carnegie Mellon University [✕]Processor Architecture Research Lab, Intel Labs [∇]Bilkent University [◇]ETH Zürich
[‡]Facebook [○]King Mongkut's University of Technology North Bangkok ^{*}University of Illinois at Urbana-Champaign

Near-memory Pre-alignment Filtering

Gagandeep Singh, **Mohammed Alser**, Damla Senol Cali, Dionysios Diamantopoulos, Juan Gomez-Luna, Henk Corporaal, Onur Mutlu,

[“FPGA-Based Near-Memory Acceleration of Modern Data-Intensive Applications”](#)

IEEE Micro, 2021.

[\[Source Code\]](#)



[Home](#) / [Magazines](#) / [IEEE Micro](#) / [2021.04](#)

IEEE Micro

FPGA-Based Near-Memory Acceleration of Modern Data-Intensive Applications

July-Aug. 2021, pp. 39-48, vol. 41

DOI Bookmark: [10.1109/MM.2021.3088396](https://doi.org/10.1109/MM.2021.3088396)

Authors

[Gagandeep Singh](#), ETH Zürich, Zürich, Switzerland

[Mohammed Alser](#), ETH Zürich, Zürich, Switzerland

[Damla Senol Cali](#), Carnegie Mellon University, Pittsburgh, PA, USA

[Dionysios Diamantopoulos](#), Zürich Lab, IBM Research Europe, Rüschlikon, Switzerland

[Juan Gomez-Luna](#), ETH Zürich, Zürich, Switzerland

[Henk Corporaal](#), Eindhoven University of Technology, Eindhoven, The Netherlands

[Onur Mutlu](#), ETH Zürich, Zürich, Switzerland

◀	▶
Previous	Next
☰	Table of Contents
📄	Past Issues

Demeter (HD Food Microbiome Profiling)

Taha Shahroodi, Mahdi Zahedi, Can Firtina, **Mohammed Alser**, Stephan Wong, Onur Mutlu, Said Hamdioui

[“Demeter: A Fast and Energy-Efficient Food Profiler using Hyperdimensional Computing in Memory”](#)

IEEE Access, 2022

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

 **RESEARCH ARTICLE**

Demeter: A Fast and Energy-Efficient Food Profiler Using Hyperdimensional Computing in Memory

**TAHA SHAHROODI^{ID1}, MAHDI ZAHEDI^{ID1}, CAN FIRTIINA², MOHAMMED ALSER^{ID2},
STEPHAN WONG¹, (Senior Member, IEEE), ONUR MUTLU^{ID2}, (Fellow, IEEE),
AND SAID HAMDIOUI^{ID1}, (Senior Member, IEEE)**

¹Q&CE Department, EEMCS Faculty, Delft University of Technology (TU Delft), 2628 CD Delft, The Netherlands

²SAFARI Research Group, D-ITET, ETH Zürich, 8092 Zürich, Switzerland

AIM (PIM Sequence Alignment Framework)

Safaa Diab, Amir Nassereldine, **Mohammed Alser**, Juan Gómez-Luna,
Onur Mutlu, Izzat El Hajj

[“A Framework for High-throughput Sequence Alignment using Real Processing-in-Memory Systems”](#)

arXiv, 2022

[\[Source code\]](#)

A Framework for High-throughput Sequence Alignment using Real Processing-in-Memory Systems

Safaa Diab¹, Amir Nassereldine¹, Mohammed Alser², Juan Gómez Luna², Onur Mutlu², Izzat El Hajj¹

¹*American University of Beirut, Lebanon* ²*ETH Zürich, Switzerland*

SneakySnake [Bioinformatics 2020]

Mohammed Alser, Taha Shahroodi, Juan-Gomez Luna, Can Alkan, and Onur Mutlu,

"SneakySnake: A Fast and Accurate Universal Genome Pre-Alignment Filter for CPUs, GPUs, and FPGAs"

Bioinformatics, 2020.

[\[Source Code\]](#)

[\[Online link at Bioinformatics Journal\]](#)

Bioinformatics



SneakySnake: a fast and accurate universal genome pre-alignment filter for CPUs, GPUs and FPGAs

Mohammed Alser ✉, Taha Shahroodi, Juan Gómez-Luna, Can Alkan ✉, Onur Mutlu ✉

Bioinformatics, btaa1015, <https://doi.org/10.1093/bioinformatics/btaa1015>

Published: 26 December 2020 **Article history** ▼

GateKeeper [Alser+, Bioinformatics 2017]

Mohammed Alser, Hasan Hassan, Hongyi Xin, Oguz Ergin, Onur Mutlu, and Can Alkan
"GateKeeper: A New Hardware Architecture for Accelerating Pre-Alignment in DNA Short Read Mapping"

Bioinformatics, [published online, May 31], 2017.

[\[Source Code\]](#)

[\[Online link at Bioinformatics Journal\]](#)

Bioinformatics

iSCB
INTERNATIONAL SOCIETY FOR
COMPUTATIONAL BIOLOGY

Article Navigation

GateKeeper: a new hardware architecture for accelerating pre-alignment in DNA short read mapping FREE

Mohammed Alser ✉, Hasan Hassan, Hongyi Xin, Oğuz Ergin, Onur Mutlu ✉, Can Alkan ✉

Bioinformatics, Volume 33, Issue 21, 01 November 2017, Pages 3355–3363,

<https://doi.org/10.1093/bioinformatics/btx342>

Published: 31 May 2017 **Article history** ▼

MAGNET

Mohammed Alser, Onur Mutlu, and Can Alkan.

["MAGNET: understanding and improving the accuracy of genome pre-alignment filtering"](#)

IPSI Transaction (2017).

[\[Source code\]](#)

MAGNET: Understanding and Improving the Accuracy of Genome Pre-Alignment Filtering

Alser, Mohammed; Mutlu, Onur; and Alkan, Can

Shouji (障子) [Alser+, Bioinformatics 2019]

Mohammed Alser, Hasan Hassan, Akash Kumar, Onur Mutlu, and Can Alkan,
"Shouji: A Fast and Efficient Pre-Alignment Filter for Sequence Alignment"
Bioinformatics, [published online, March 28], 2019.

[\[Source Code\]](#)

[\[Online link at Bioinformatics Journal\]](#)

Bioinformatics, 2019, 1–9

doi: 10.1093/bioinformatics/btz234

Advance Access Publication Date: 28 March 2019

Original Paper



Sequence alignment

Shouji: a fast and efficient pre-alignment filter for sequence alignment

**Mohammed Alser^{1,2,3,*}, Hasan Hassan¹, Akash Kumar², Onur Mutlu^{1,3,*}
and Can Alkan^{3,*}**

¹Computer Science Department, ETH Zürich, Zürich 8092, Switzerland, ²Chair for Processor Design, Center For Advancing Electronics Dresden, Institute of Computer Engineering, Technische Universität Dresden, 01062 Dresden, Germany and ³Computer Engineering Department, Bilkent University, 06800 Ankara, Turkey

In-Memory Sequence Analysis GRIM-Filter

- Jeremie S. Kim, Damla Senol Cali, Hongyi Xin, Donghyuk Lee, Saugata Ghose, **Mohammed Alser**, Hasan Hassan, Oguz Ergin, Can Alkan, and Onur Mutlu, "**GRIM-Filter: Fast Seed Location Filtering in DNA Read Mapping Using Processing-in-Memory Technologies**"
to appear in [BMC Genomics](#), 2018.
*Proceedings of the [16th Asia Pacific Bioinformatics Conference \(APBC\)](#),
Yokohama, Japan, January 2018.*
[arxiv.org Version \(pdf\)](#)

BMC Genomics

Research | [Open Access](#) | [Published: 09 May 2018](#)

GRIM-Filter: Fast seed location filtering in DNA read mapping using processing-in-memory technologies

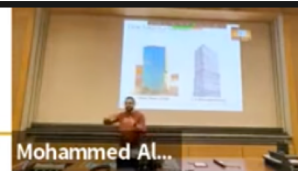
[Jeremie S. Kim](#) ✉, [Damla Senol Cali](#), [Hongyi Xin](#), [Donghyuk Lee](#), [Saugata Ghose](#), [Mohammed Alser](#),
[Hasan Hassan](#), [Oguz Ergin](#), [Can Alkan](#) ✉ & [Onur Mutlu](#) ✉

[BMC Genomics](#) **19**, Article number: 89 (2018) | [Cite this article](#)

4340 Accesses | **39** Citations | **9** Altmetric | [Metrics](#)

Accelerating Genome Analysis

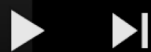
How Large is a Genome?



Prime Tower, Zurich



~3.2 billion genomic bases



7:02 / 1:44:35

SAFARI

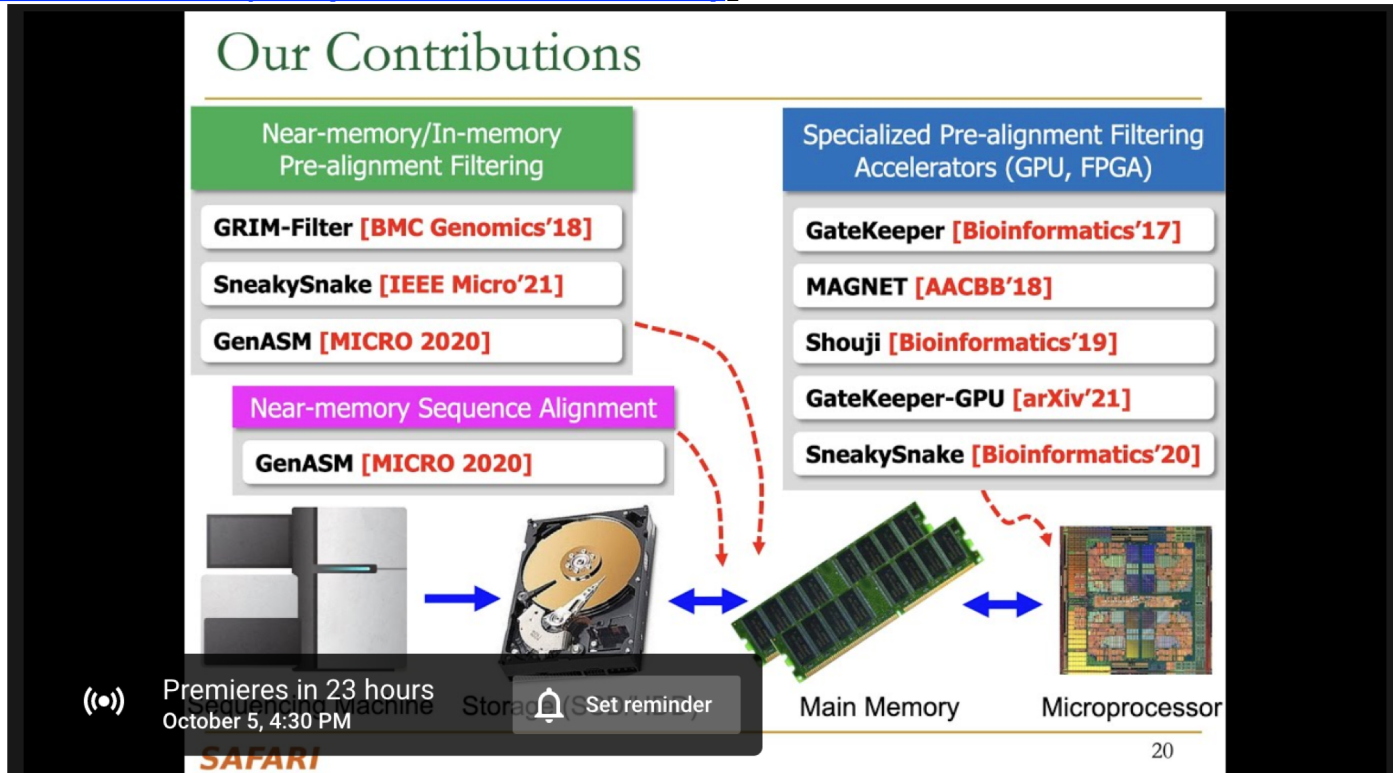


Livestream - Seminar in Computer Architecture - ETH Zürich (Spring 2022)

Seminar in Computer Arch. - Lecture 5: Accelerating Genome Analysis (Spring 2022)

More on Accelerating Genome Analysis ...

- Mohammed Alser,
"Accelerating Genome Analysis: A Primer on an Ongoing Journey"
Talk at [RECOMB 2021](#), Virtual, August 30, 2021.
[[Slides \(pptx\)](#) ([pdf](#))]
[[Talk Video](#) (27 minutes)]
[[Related Invited Paper](#) (at [IEEE Micro](#), 2020)]



More on Intelligent Genome Analysis ...

- Mohammed Alser,
"Computer Architecture - Lecture 10: Intelligent Genome Analysis"
ETH Zurich, Computer Architecture Course, Fall2021, Lecture 10, Virtual, 29 October 2021.
[[Slides \(pptx\)](#)] [[pdf](#)]
[[Talk Video](#) (3 hour 2 minutes, including Q&A)]
[[Related Invited Paper \(at IEEE Micro, 2020\)](#)]



Computer Architecture - Lecture 10: Intelligent Genome Analysis (Fall 2021)

412 views • Streamed live on Oct 29, 2021

👍 19 🗨️ 0 ➦ SHARE ≡+ SAVE ...

More on Intelligent Genome Analysis ...

- Mohammed Alser,
"Computer Architecture - Lecture 8: Intelligent Genome Analysis"
ETH Zurich, Computer Architecture Course, Lecture 8, Virtual, 15 October 2021.
[[Slides \(pptx\)](#)] [[pdf](#)]
[[Talk Video](#) (2 hour 54 minutes, including Q&A)]
[[Related Invited Paper \(at IEEE Micro, 2020\)](#)]

Our Solution: GateKeeper

Alignment Filter + FPGA-based Alignment Filter = 1st FPGA-based Alignment Filter.

Low Speed & High Accuracy
Medium Speed, Medium Accuracy
High Speed, Low Accuracy

$\times 10^{12}$ mappings

$\times 10^3$ mappings

1 High throughput DNA sequencing (HTS) technologies
2 Read Pre-Alignment Filtering Fast & Low False Positive Rate
3 Read Alignment Slow & Zero False Positives

108

SAFARI

2:08:58 / 2:54:18 • GateKeeper >

ETH ZENTRUM

Computer Architecture - Lecture 8: Intelligent Genome Analysis (ETH Zürich, Fall 2020)

More on Fast Genome Analysis ...

- Onur Mutlu,
"Accelerating Genome Analysis: A Primer on an Ongoing Journey"
Invited Lecture at [Technion](#), Virtual, 26 January 2021.
[[Slides \(pptx\)](#)] [[pdf](#)]
[[Talk Video](#) (1 hour 37 minutes, including Q&A)]
[[Related Invited Paper \(at IEEE Micro, 2020\)](#)]

Insight: Shifting a String Helps Similarity Search

7 matches 1 mismatch

ISTANBUL

ISTNBUL

ISTNBUL

81

Onur Mutlu - Invited Lecture @Technion: Accelerating Genome Analysis: A Primer on an Ongoing Journey

566 views · Premiered Feb 6, 2021

👍 31 💬 0 ➦ SHARE ≡ SAVE ...

 Onur Mutlu Lectures
13.9K subscribers

ANALYTICS EDIT VIDEO

Detailed Lectures on Genome Analysis

- **Computer Architecture, Fall 2020, Lecture 3a**
 - **Introduction to Genome Sequence Analysis** (ETH Zürich, Fall 2020)
 - <https://www.youtube.com/watch?v=CrRb32v7SJc&list=PL5Q2soXY2Zi9xidyIgBxUz7xRPS-wisBN&index=5>
- **Computer Architecture, Fall 2020, Lecture 8**
 - **Intelligent Genome Analysis** (ETH Zürich, Fall 2020)
 - <https://www.youtube.com/watch?v=ygmQpdDTL7o&list=PL5Q2soXY2Zi9xidyIgBxUz7xRPS-wisBN&index=14>
- **Computer Architecture, Fall 2020, Lecture 9a**
 - **GenASM: Approx. String Matching Accelerator** (ETH Zürich, Fall 2020)
 - <https://www.youtube.com/watch?v=XoLpzmN-Pas&list=PL5Q2soXY2Zi9xidyIgBxUz7xRPS-wisBN&index=15>
- **Accelerating Genomics Project Course, Fall 2020, Lecture 1**
 - **Accelerating Genomics** (ETH Zürich, Fall 2020)
 - <https://www.youtube.com/watch?v=rgjl8ZyLsAg&list=PL5Q2soXY2Zi9E2bBVAgCqLgwiDRQDTyId>

Prior Research on Genome Analysis (1/2)

- Alser+, "[Technology dictates algorithms: Recent developments in read alignment](#)", *Genome Biology*, 2021.
- Alser + "[SneakySnake: A Fast and Accurate Universal Genome Pre-Alignment Filter for CPUs, GPUs, and FPGAs.](#)", *Bioinformatics*, 2020.
- Senol Cali+, "[GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis](#)", *MICRO* 2020.
- Kim+, "[AirLift: A Fast and Comprehensive Technique for Translating Alignments between Reference Genomes](#)", *arXiv*, 2020
- Alser+, "[Accelerating Genome Analysis: A Primer on an Ongoing Journey](#)", *IEEE Micro*, 2020.

Prior Research on Genome Analysis (2/2)

- Firtina+, "[Apollo: a sequencing-technology-independent, scalable and accurate assembly polishing algorithm](#)", *Bioinformatics*, 2019.
- Alser+, "[Shouji: a fast and efficient pre-alignment filter for sequence alignment](#)", *Bioinformatics* 2019.
- Kim+, "[GRIM-Filter: Fast Seed Location Filtering in DNA Read Mapping Using Processing-in-Memory Technologies](#)", *BMC Genomics*, 2018.
- Alser+, "[GateKeeper: A New Hardware Architecture for Accelerating Pre-Alignment in DNA Short Read Mapping](#)", *Bioinformatics*, 2017.
- Alser+, "[MAGNET: understanding and improving the accuracy of genome pre-alignment filtering](#)", *IPSI Transaction*, 2017.

P&S Accelerating Genomics

Introduction & Course Logistics

Dr. Mohammed Alser

 @mealsr

ETH Zurich

Fall 2022

27 October 2022

SAFARI

ETH zürich