CDC HEALTH DATA INNOVATION SUMMIT

Atlanta, GA | September 27-28, 2023







Kelsey Florek, PhD, MPH Senior Genomics and Data Scientist Wisconsin State Laboratory of Hygiene

linkedin.com/in/kelseyflorek

Evan Davey, BCS Partner & Global Lead DEPT®/HEALTH

linkedin.com/in/evandavey

DEPT_®



Today

- ⁰¹ Introduction
- ⁰² Why Genomic Sequencing?
- ⁰³ What is Easy Genomics?
- 04 AWS HealthOmics
- 05 Demonstration

Introduction





Evan Davey, BCS Partner & Global Lead DEPT®/HEALTH

linkedin.com/in/evandavey

02

Why Genomic Sequencing?

Easy Genomics: Democratising Access to Genomic Sequencing

What is Genomic Sequencing?

- → All organisms carry a genome or blueprint composed of a sequence of nucleotides (A, C, G, and T/U) that are used to construct proteins that support a variety of functions.
- → Sequencing an organism's genome is a powerful step in understanding the potential characteristics or features an organism might have.
- → Genomics enhances public health outcomes by tailoring healthcare interventions, improving disease surveillance, and advancing our understanding of genetic factors affecting health and disease.



Why is it important?

- → Genomic Sequencing contributes to disease prevention, personalized healthcare, disease tracking, and the development of effective treatments.
- → During the Covid-19 pandemic Genomics played a critical role in tracking variants across populations and geographies



Revolutionary Technology Next Generation Sequencing

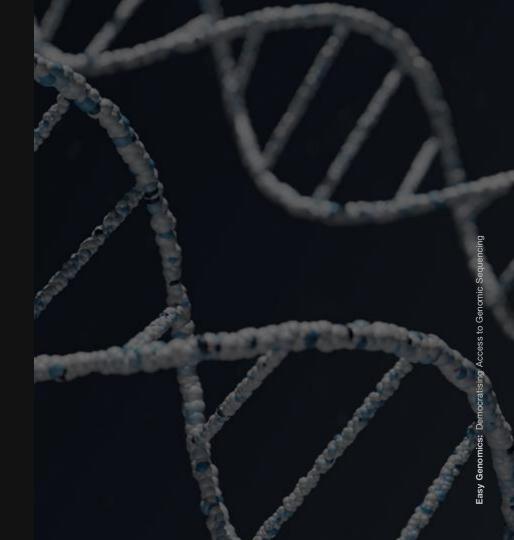
- Next Generation Sequencing
- \rightarrow massively parallel approach to sequencing a genome
- \rightarrow 360 billion ATGC's per sequencing run
- \rightarrow 40,000 150,000 words in a novel
- \rightarrow average word length in English is 4.79
- → one 72 hour sequencing run would generate 791,121 novels with 95,000 words each

Revolutionary Challenges

- → Cost Genomic sequencing can be expensive, making it inaccessible or challenging for areas of the healthcare system less resourced
- → Training Healthcare professionals and public health workers require additional training to effectively use genomic sequencing data in their practice.

03

What is Easy Genomics?



🔁 Easy Genomics

Open source for all.

Easy Genomics is a no-code, open-source solution for genomic data analysis. If the past few years have taught us anything, it's that genomic sequencing will play a major part in the future of (public) health. But it's an intensive, complex and expensive process to set up and run.

We've changed that.

</>>

Simple to use.

Requires little to no technical expertise.

\odot

Seamless experience

Powered by NextFlow, we provide a faster, hassle-free workflow.

Ĥ

Secure data storage

Endless scalability as your needs grow.

dt.

Start small, scale fast

Encrypted security to keep your patient data locked and secure.

Easy Genomics

How it works

Run your genomic workload on a no-code front-end interface, that any nontechnical user can operate in three easy steps:

Select a Workflow

0

ø

Customise to your analytic requirements.

Upload your Data

Manage access to single & multiple samples.

Hit Submit

Monitor & receive your results via browser or email.

dSb5dOgs2vmgi.eloudfront.net/workflow/launch		6
0 EasyGenomics	Runs	Work Rows Users My Profile
Workflows		
List of all of available workflows.		Add workflow
Lab Name 1 Lab Name 2		
Name	Cescription	Туре
gata-garmfine-slips-indels (1-1746/301746)4(1+)	Well-Road for perifficial short variant sharvery using GADOC in NextRoad	Ready to But Workfree
nf-core-ampliang IC Salartonication	102.455A angleton sequencing scaleto evolution asing QMEL	Privato Wateflow
M-core-attacked 41-24/14/11/11/22/24	ATHC was pash rating. Of and differential analysis position	(Picate Workford)
N ⁴ -core-becase 30-24412/met247342	freque taccurist assumption prof assumption population.	Private Workfow
nf-core-chipseq kp-thoosenssisseps	Drift-uniq peak-calling, DC and differential analysis applies.	Private Wetches
nt-core-mag 46-tech/listiticat/77	Researchery and Schweig of Instrugeneerses.	Private Workflow
nf-core-mhoguant	startify and quartify RHC status position base	Ready to Hart Machines

Easy Genomics



Easy Genomics: Democratising Access to

Easy Genomics V1

- \rightarrow User accounts and authentication
- → Simple process for uploading and/or submitting genomic sequencing data for analysis
- \rightarrow Allow users to select and easily configure analyses
- \rightarrow Access to analytical workflow status and results



Easy Genomics V1

- \rightarrow 3 user types: super admin, lab admin, user
- → Supports multiple "labs" with separate workflows and users
- → Enables data submission through web interface into AWS S3 and automated samplesheet generation
- → Runs analytical workflows through Nextflow Tower API

© Karş Ownamit L			Apra Monthan A	
Launch wo	orkflow			
Mage 1 And San Januar	Bag 2 Ignat Departor	Bree 3 (Last Parameters)	Fing 4 Annual Matthew	
Capitual DNA Sequence for			Latinat film	
Sampin D WE WILL CITIZE	Pair 1 Wirefour 202085,01 Serie pr	Progress Pair 2	Program • 1223MLA2 hers git	



AWS Health Omics



AWS Health Omics



AWS Health Omics

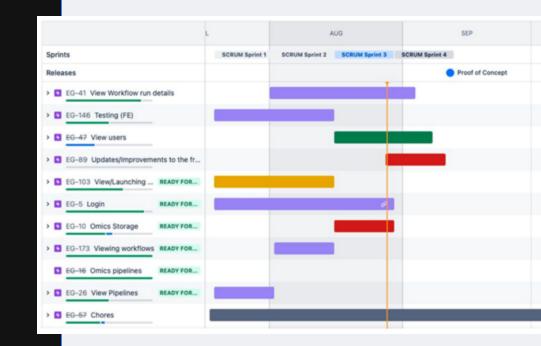
- → Removes the need to configure and manage complex cloud infrastructure involving AWS Batch / EC2
- → Provides tailored storage for sequence and variant data
- \rightarrow Supports Nextflow & WDL workflows
- \rightarrow Requires staging workflow containers and data
- \rightarrow Challenging workflow configuration
- \rightarrow Accessible through API



Now we have Health Omics, let's upgrade Easy Genomics!

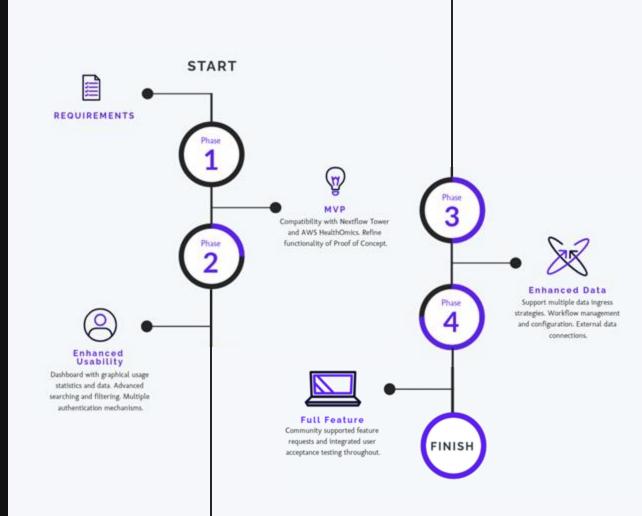
- → Develop an open sourced solution powered by AWS' HealthOmics
- → Provide an easy-to-use, web-based interface that will facilitate access to key genomic sequencing pipelines for all lab technicians
- \rightarrow With limited training and no technical skills

Easy Genomics v2



Easy Genomics v2 Roadmap

- \rightarrow August 2023 POC Launched
- → Q1 2024 Initial Release



Open source + community driven

\rightarrow Open Source

Easy Genomics is built for a field that is fundamentally built on open source projects. Genomic software and workflows are community efforts to improve our understanding of the biological world that we exist in.

\rightarrow Community Driven

Easy Genomics is a solution designed for a community of public health laboratories and is charged with meeting their unique needs.

Launch wo	orkflow			
Dep 1	Tree 2	(mg)	Transa (Sec.	
hat this therein.	Upinal Separce	bat Parameters	Barriese Marielland	
Upload ONA Seguence N			agreed from	
C feet amount not unline				
Bangle (D	Par 1	Progress Par 2	Program	
WF-W3LH-222388	www.witzwitztatt,#thung.gr	www.ueu	Cash, Ko Away, gal	

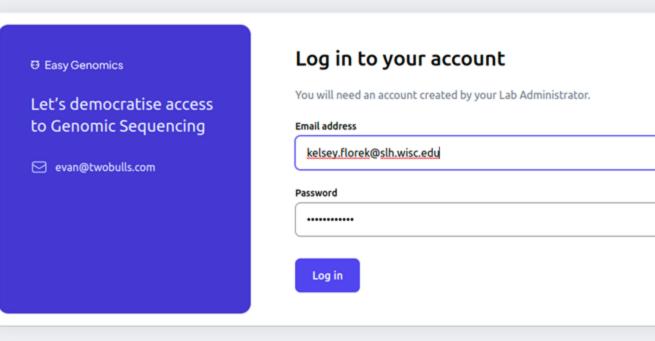
05

Demonstratio

n

Easy Genomics: Democratising Access to Genomic Sequencing

🔁 Easy Genomics



••••]

••••]

Orkflows	aunch a pipeline and start testing.		Launch pipeline
Pipeline	User	Activity	
wi-M5192-230818_4 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	SUCCEEDED	
WI-M5192-230818_3 wslh-blo/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	
WI-M5192-230829_4 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	SUCCEEDED	
wi-M5192-230829_3 wsih-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	
WI-M5192-230829_2 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	
WI-M5192-230818_2 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	
WI-M5192-230829 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	
WI-M3478-230825 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	SUCCEEDED	
Wi-M3478-230821 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	SUCCEEDED	
WI-M5192-230818 wslh-bio/spriggan	dustin-lyfoung Submitted: 4 days ago	FAILED	

🔁 Easy Genomics

Easy Genomics			Workflows	Sequencing	Pipelines	Users	KF
Sequencing Upload DNA Sequencing Reads Read upload instructions					Sample sheet	Upload	files
Sample ID	Pair 1	Progress	Pair 2			Progr	ress
TestSample01-211229_S1_L001	TestSample01-211229_S1_L001_R1_001.fastq.gz	100%	TestSample01-21	1229_S1_L001_R2	_001.fastq.gz	100	%

	Sequencing				and the	-			
	Upload DNA Sequencing Reads				Sample sheet	Upload files			
	Read upload instructions								
	Sample ID	Pair 1	Progress	Pair 2		Progress			
	TestSample01-211229_51_L001	TestSample01-211229_51_L001_R1_001.fastq.gz	100%	TestSample01-211229_51_L	001_R2_001.fastq.gz	100%			
		sample sheet.crv	LibraOffice Cale					- 0	
le Edit View Insert F	ormat Styles Sheet Data Tools Window		Contraction Care						×
		-∂- <u>@</u> * ≣- ∎-149434 ⊽			.				
		E ∃ <u>5</u> ≣ ∓ ‡ ≜ 💁 •% 00 D	1 .04 .0X 🖄	≝ ⊞• ⊪• ⊻.	- B - B				
- fx Σ	• = sample							•	
A sample	and the second s	0			c		D		
TestSample01-211229_51	100 sl. itev esh sequencing inboxics4553d9-21d0-	4d03-94a3-6163763b7700TestSample01-211229_51_L001_R1_0	fastg_2 01+s3.idex.esib.seguer	ncing_inboxicx4653d9-21d0-4d03	-94a3-6163763b7708Test5	ample01-211229_51_L0	01_R2_001.8xstq.gz		
TestSample01-211229_51_	page_1 construction_add_area_add_area_add_area_add_area_add_area_add_area_add_area_add_area_add_area_add_area_add_area	4d03-04a3-6163763b7700TestSample01-211229_51_L001_R1_0	fang, 2 11×13 Her with respec	ncing_integratic4653d9-21d0-4d03	94a3-6163763b770#Test5	ample01-211229_51_L0	01_R2_001.fastq.gz		A
TestSample01-211229_51_	1990, 1 OF 11 igns with sequencing indexicologistic 455349-2140-	4303-94a3-636376357708TeutSample01-211229_51_L001_81_0	fanta, 2 11-s3 itter estit seque	ning (régyik:465349-2140-4403	94a3-6163763b7708Test5	ample01-211229_51_L0	01_R2_001.twitq.gz		Æ
Tex5ample01-211229_51	Den 1 inger weit sewencing obsvicz465349-2140-	4003-94a3-6363763b7700TextSample01-211229_51_1001_R1_0	fang, 2 It-strides wolft in gen	ncing_inboskc4553d9-2140-4401	-94a3-6163763b7708Test5	ample01-211229_51_L0	01_R2_001 fixitig gr		A
TestSample05-211229_51	Ling in the second product of the second pro	4603-94a3-6363763b7709TextSample03-213229_51_L001_R1_0	1909, 2 11-51-997, 958, 958, 982, 99	nong-jobgsko-4653d9-2140-4401	94a3-6163763b770fText5	ampie01-211229_51_L0	01_R2_001.fkstq.gz		Æ
TestSample01-211229_51	and a second sec	4603-94a3-61637630:7709TextSample01-211229_51_L001_R1_0	1992, 2 11-11-999-900-99249	10009-10804053149-2140-4403	5-94a3-6163763b7708Teud5	ample01-211229_51_L0	01_R2_0018x89.gz		A 2
TestSample05-211229_51_	2000 - 2000 - 2000 - 2000 - 2000 - 2000 - 2100 -	4603-94a3-4183763b7709TextSample01-211229_51_L001_91_6	fang, 2 11-så iden mölt inspire	9 019-109 9455365-2140-4401	5-94a3-616376367708Tesz5	ample01-211229_51_L0	01_R2_001 fixetq.gr		A
Tep2Sample01-211229_51_ 	00 ⁰⁰⁰ 000 <u>0000000000000000000000000000</u>	4603-94a3-4383763b7709TextSample01-213229_51_L001_91_0	949,2 11-12-1992-490-549241	nong-inbyetceks349-2140-4401	-94a3-6163763b7708Teut5	ample01-211229_51_L0	01_R2_001 fixetq.gr		A 2
Tendiampie01-211229_51_		4603-94a3-6383763b7709TextSample03-213229_51_L001_R1_0	900, 2 11 - 12 - 1992, and 2 - 2002	5079-1000/to485369-2340-4401	1-84a3-616376367708Teut5	ampim01-211229_51_L0	01_R2_001.faxing.gz		A 2
2 TestSampte02-211229_51_ 5 5 6 7 7 8 8 9 9 9 9 9 9 9 8 8 5 6	2000 1990 - 19900 - 19900 - 19900 - 1990 - 19900 - 1990 - 1990 - 1990 - 1990 -	4603-94a3-4183763b7709TestSample03-211229_51_L001_91_6	900, 2 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	<u>nang idoo</u> tic465369-2140-4401	-94a3-61637630-770fTead5	angie01-211229_51_L0	01_R2_001.faxing.gz		A 2
TentSampar03-311229_51	2007.000 (0000000000000000000000000000000	4603-94a3-4383763b7709TextSample03-233229_51_L001_91_6	908,2 Li-sl-1990,400,400	1000 indepite465345-2140-4401	194a3-6183763b77087ead5	ampi#01-211229_51_L0	01_82_001.http.gr		A 2
TentSampse0: 211229 81 S O O Z S O S O S O S O S	COn Configure and accounting indepictories 21:00-	4603-94a3-4383763b7709TextSample01-233229_51_L001_91_0	1900,2 11-51 / 900,900,900,900	<u>nong kon</u> vice485349-2140-4401	-94a3-6183763b77087ead5	ampi#01-211229_51_L0	01_R2_001 http:		A 2
TestSampse0: 211229_51 S	2000 1999 - 2000 - 2000 - 2000 - 2000 - 2000 - 2100 -	4603-94a3-4183763b;7709TestSample03-211229_51_L001_92_6	990.2 19 - M - Sec. and Longer	<u>nang idoo</u> tic465369-2140-4401	194a3-61637635-770/Teat5	ampie01-211229_51_10	01_92_001 honty gr		A 2
TentSampte03-311229_51 S O	2000 (1996) - mail ang	4603-9463-4183763b7709TextSample03-211229_51_L001_91_6	1990.2	<u>nong iden</u> eto465349-2140-4401	194a3-61637630-770fTeat5	ample01-211229_51_00	01_R2_001 holty gz		A 2
TentSampse0: 211229 %1 to		4603-94a3-4383763b7709TextSample03-233229_51_L001_91_6	in the second second	1010 idoreito405345-2140-4401	194a3-6183763b7708Teat5	angle01-211228_51_L0	01_92_001 Bostq gz		A 2
VerdSampiet2-211229_51_		4603-94a3-4183763b;7709TestSample03-211229_51_L001_92_6	1. 900.2 1 U. Spr. and sec.er	<u>aang idoo</u> tice465369-2140-4401	194a3-61637635-770/Teat5	angle01-211228_51_L0	01_R2_001 Bonty gz		A 2
TereSampse0: 211229_51		4603-9463-4183763b:7709TextSample03-211229_51_0001_91_6	12-900.2	<u>nong iden</u> tic465349-2140-4401	194a3-61637635-770fTeat5	ampie01-211229_51_L0	01_R2_001 Bonty gz		A 2
TentSampse0: 211229 11 S		4603-9463-4183763b7709TextSample03-211229_51_L001_91_6	10 900.2	nong inderetice465345-2140-4401	194a3-61637635-770fTeat5	angle01211228_51_60	01_R2_001 honty gz		A 2
TendSampse0: 211229 81 S		4603-94a3-4183763b;7709TestSample03-211229_51_L001_92_6	11- 900,2	<u>nong ido e</u> tc465369-2140-4401	194a3-61637635770/Teat5	angle01-211228_51_L0	01_R2_001 hotty gz		A 2
TestSample0: 211229 51 S Control of the second seco		4603-9463-4183763b:7709TextSample03-211229_51_0001_92_6	11- 900,2	<u>nong ido e</u> tc445349-2140-4401	194a3-61637635-770/Teat5	argie01-211229_51_L0	01_R2_001 hosts gz		A 2
TentSampse0: 211229 81 L Control of the second sec		4603-9463-4183763b7709TextSample03-211229_51_0001_91_6	10- 300,2	<u>nong iden</u> eto465349-2140-4401	194a3-61637635-770fTeat5	argie01211228_51_60	01_R2_001 holty gz		A 2
TentSampset2 311229 81 S		4603-94a3-4183763b;7709TestSample03-211229_51_L001_92_6	11- 300. 2	<u>nong ido e</u> tc446349-2140-4401	194a3-61637635770/Teat5	argie01-211228_51_L0	01_R2_001 hotty gz		A 2
2 TestSampleO 211229 51 5 5 6 7 8 9 9 9 9 9 9 9 9 9 9 9 9 9		4603-9463-4183763b:7709TextSample03-211229_51_0001_92_6	12- 300. 2	<u>nong i do e</u> lco465369-2140-4401	1.94a3-61637635-770/Teat5	argie01-211229_51_L0	01_R2_001 Bonty gz		A 2
2 TestSampiet2-211229_51 4 5 6 7 7 8 9 9 9 9 9 9 9 9 9 9 9 9 9		4603-9463-4183763b;7709TextSample03-211229_51_0001_92_3	10-900-2	nong.idoutce465345-2140-4401	194a3-61637635-770fTeat5	argie01-211228_51_L0	01_R2_001 hosts gz		A 2
2 TestSampset2 311229 51 5 5 5 5 5 5 5 5 5 5 5 5 5		4603-9463-41837636:7709TestSample03-211229_51_(001_96_6	11- 300. 2	<u>nong ido e</u> tce465369-2140-4401	194a3-61637635770/Teat5	argie01-211228_51_L0	01_R2_001 Both pr		A 2

Easy Genomics

easygenomics.org

DEPT/HEALTH



deptagency.com/health



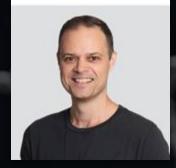


Kelsey Florek, PhD, MPH

Senior Genomics and Data Scientist Wisconsin State Laboratory of Hygiene

linkedin.com/in/kelseyflorek





Evan Davey, BCS Partner & Global Lead DEPT®/HEALTH

linkedin.com/in/evandavey

Thank you.