

DELL

The disruptive nature of portable sequencing

Nvidia GTC, 12th - 16th April 2021

Dr Miles Benton

Human Genomics
Institute of Environmental Science and Research (ESR)
Aotearoa (New Zealand)



Advocates for reproducible research...

Where possible all presentations and code are available online



github
SOCIAL CODING



sirselim.github.io/presentations



Who are we?

You may have seen recently...

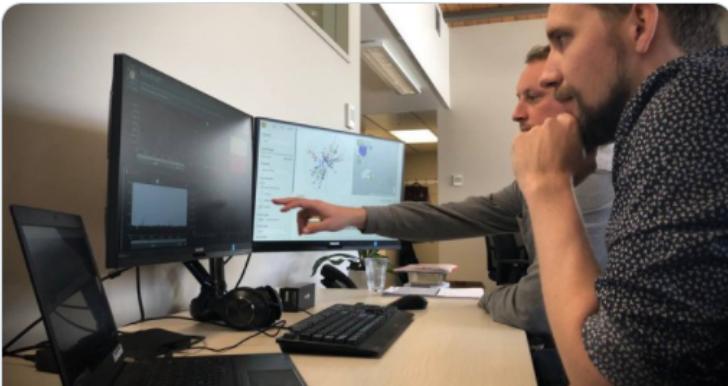
Genomics in NZ's pandemic response

A sequencing success story made possible by the awesome **Una Ren, Matt Storey & Joep de Ligt**



Jamie Morton @Jamienzherald · Mar 8

If you're one of the clueless people bashing scientists over not knowing enough about Covid-19, you may be interested to learn they're now using positive samples to rapidly sequence whole freaking genomes, right here in NZ



The Mars Bar-sized gadget helping NZ scientists decode Covid-19 virus
NZ scientists are now sequencing whole genomes of the virus driving the Covid-19 crisis.

nzherald.co.nz

Give us some time and money and we'll demonstrate what this technology can do.



NZ ChiefSciAdvisor ✅ @ChiefSciAdvisor · 13h

Whole genome sequencing has made a massive difference to containing the Auckland outbreak. Mike Bunce, CSA @EPAgovtNZ has done a rapid review for @minhealthnz CSA @ianTown3 on this great work from @ESRNewZealand to inform future outbreaks. Read it here



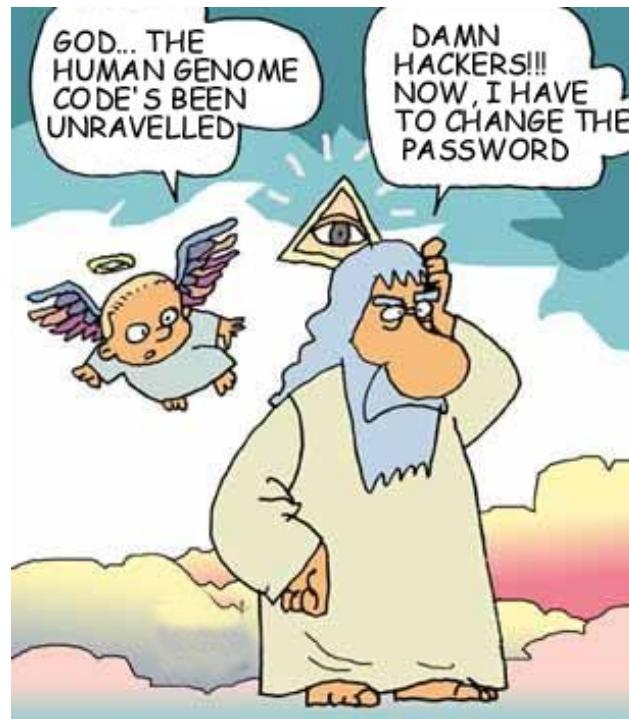
Using genetic fingerprints in our COVID-19 response
We reflect on the role of genome sequencing in the most recent COVID-19 outbreak in Auckland as the ...
pmcsa.ac.nz

Genomic sequencing

A snapshot of the Human Genomics group

- **Independent research**
 - type-2 diabetes & obesity (genomics, epigenetics, protein 'omics)
 - Guinea pigs! ([Max Berry, University of Otago](#))
 - empowering people to explore and interpret their data
- **Clinical diagnostics through genetics**
 - Contract(s) with CCDHB [Wellington Regional Genetics Lab](#)
 - Providing support to CCDHB for Clinical Genomics GA project ([Stephen Robertson](#))
- **Clinical research**, in collaboration with local clinician ([Richard Carroll](#)) - hyperparathyroidism
 - sequencing 100 exomes
 - engaging clinicians from other DHBs
 - exploring the replacement of gene panels with exome/genome

A big part of what we do is sequencing

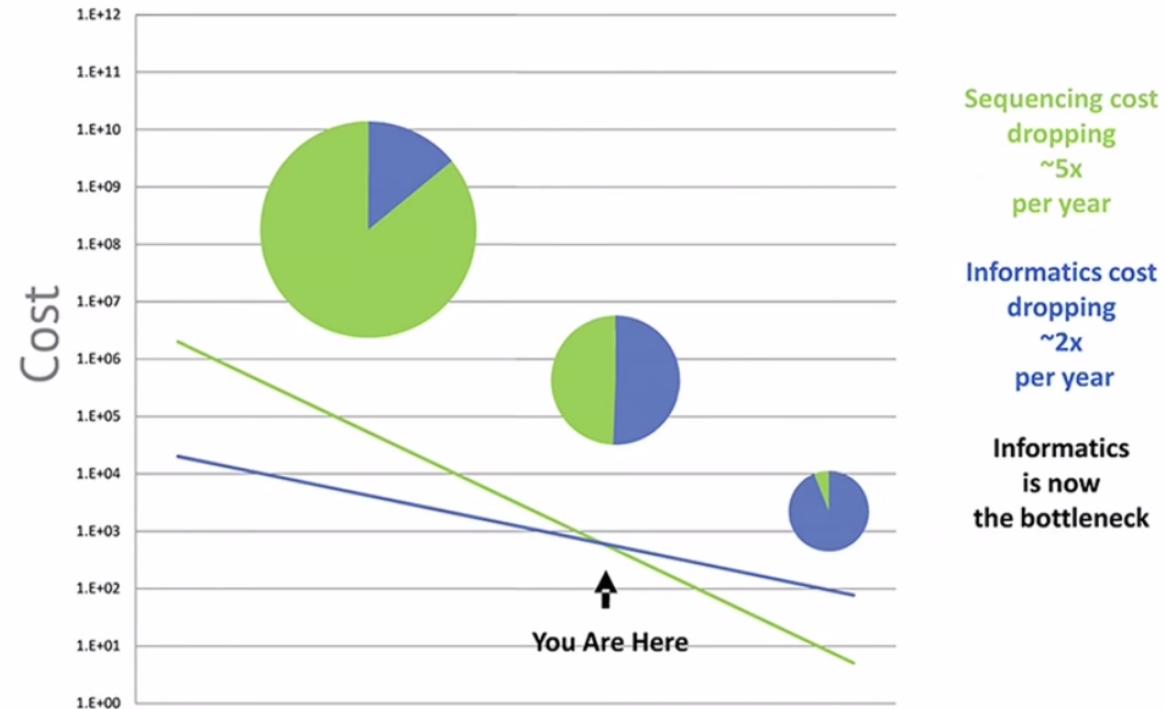




(image source: illumina.com)

"Soon, It Will Cost Less To Sequence A Genome Than To Flush A Toilet -- And That Will Change Medicine Forever"

DNA Sequencing Economics



(image source: <https://www.businessinsider.com.au/super-cheap-genome-sequencing>)

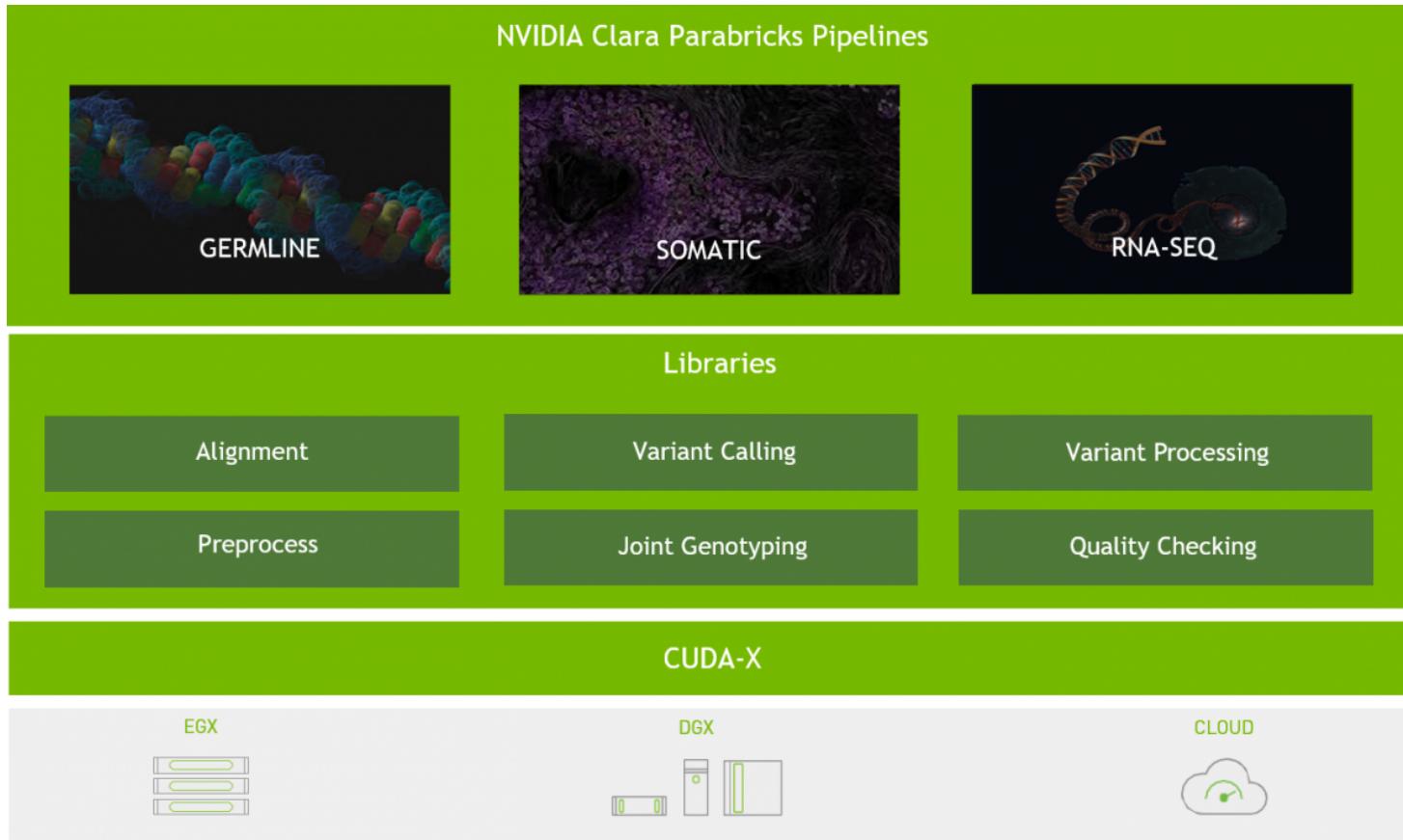
Can we speed things up?

Thanks to eResearch 2020 ...enter [Nvidia Clara Parabricks](#)



[powered by GPU](#)





www.nvidia.com/en-us/healthcare/clara-parabricks/

How much does that matter?

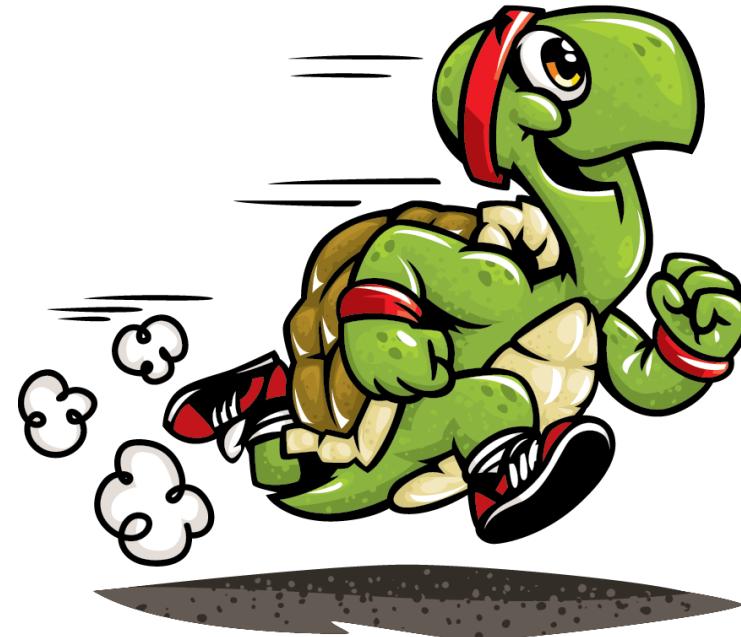
Identical pipelines with a slight difference...

- Pipeline [CPU]

- Exomes: >100X exome takes between 8-12 hours
- Genomes: ~30X genome takes between 30-48 hours

- Pipeline [GPU - Parabricks]

- Exomes: >100X exome takes between **6-12 mins**
- Genomes: 100X genome takes **~4 hours**



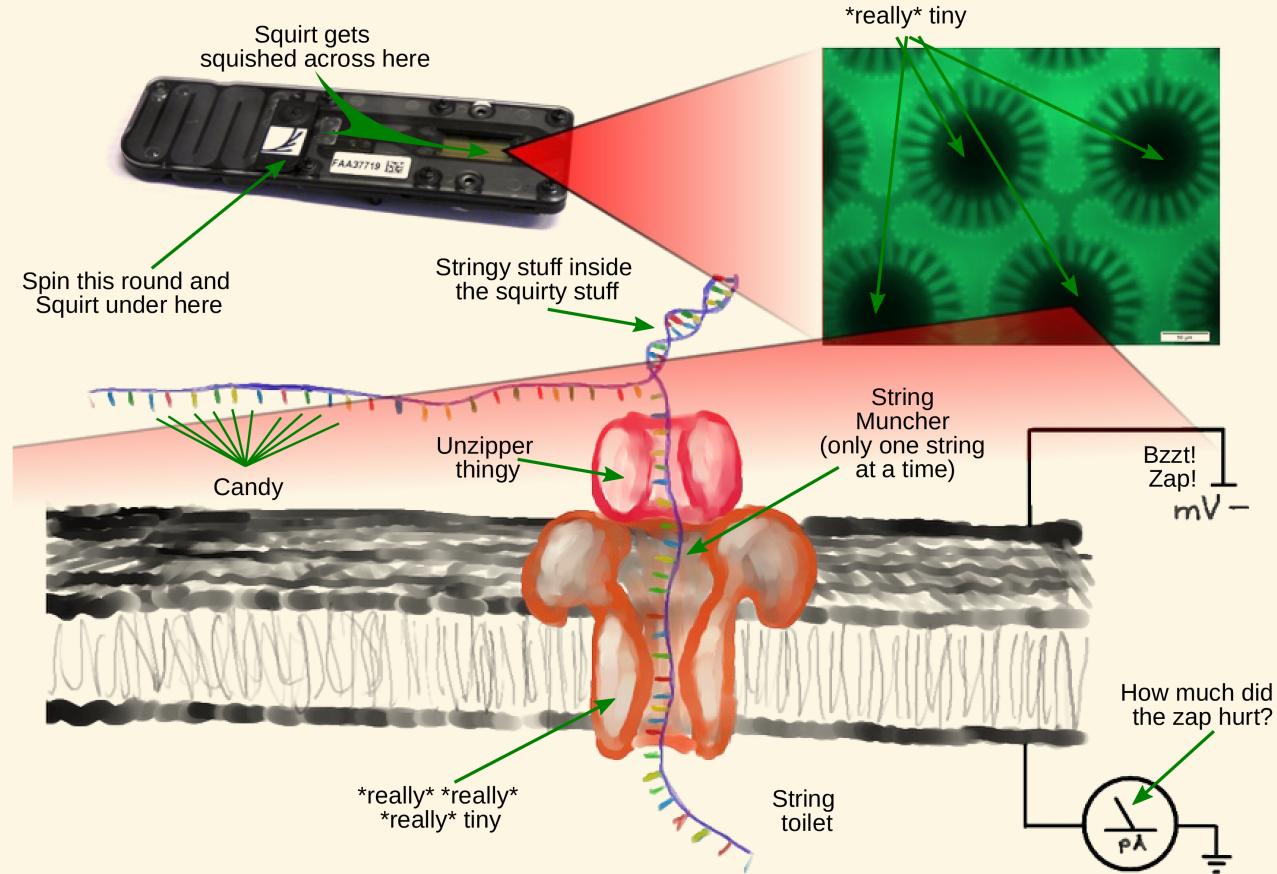
Another type of sequencing technology

...sorry PacBio I'm ignoring you today...



(image source: [nanoporetech](#))

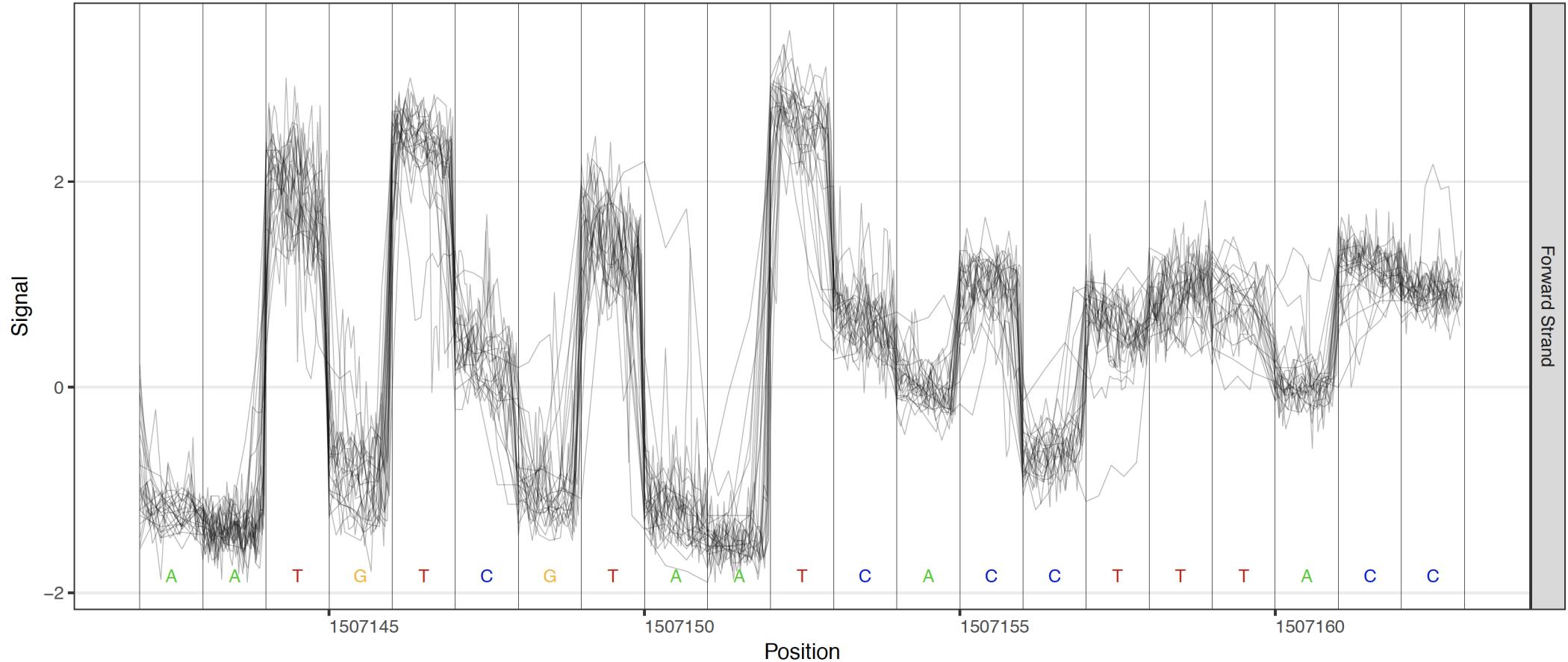
Nanopore Sequencing



How nanopore sequencing works



Example squiggle plot



(image source: [tombo manual](#))

GPUs accelerate basecalling of Nanopore long-read data (a lot)



+



=



Nanopore basecalling benchmarking: [link](#) (ESR GitHub)

Why GPUs?

New type of sequencing data requires new type of 'analysis'

- squiggle data lends itself nicely to neural nets
- GPUs are very capable in this space
 - CUDA cores

ESR GPU basecalling benchmarks ([link](#))

- Titan RTX & 2x Tesla V100



[UPDATE:] guppy is now able to scale across multiple GPUs!

Let's step back a couple of years

... a personal story ...



a personal story

- had a lumbar puncture ('9 days' old)
 - (after being born 8 weeks early)
- suspected meningitis (bacteria/viral)?
 - told after the fact
- **>48 hours** for results return

a personal story

... imagine a more ‘real-time’ procedure:

- a sample (i.e. blood) taken,
- DNA/RNA extracted/enriched,
- sequencing library prepared,
- nanopore sequencing initiated,
- sequence starts coming in almost instantly,
- **time to first read-match detection <1 hr**

This isn't a stretch of the imagination anymore, soon to be reality

Standing by for live transmission...

... live video content ...

what are we doing?

Portable 'real-time' sequencing for the masses?

- **the idea**

- low cost (< many mobile phones)
- accessible
- portable
- robust (...may have "performed" a drop test...)
- fun!

- **example use cases**

- field sequencing (real-time monitoring, forensics, agriculture, waterways, ...)
- **clinical settings**
- community outreach / teaching

Community Science | "United in Data"

Our experiences with the Xavier (link)



```
NVIDIA Jetson AGX Xavier - Jetpack 4.2.2 [L4T 32.2.1]
CPU1 [||||||] Schedutil - 12% 2.3GHz CPU5 [|||] Schedutil - 4% 2.3GHz
CPU2 [||||||] Schedutil - 12% 2.3GHz CPU6 [||||] Schedutil - 7% 2.3GHz
CPU3 [|||||||] Schedutil - 18% 2.3GHz CPU7 [||||||] Schedutil - 16% 2.3GHz
CPU4 [||||||] Schedutil - 11% 2.3GHz CPU8 [|||||||||||||] Schedutil - 44% 2.3GHz

NTS FG [ 0% ] BG [ 0% ]
Mem [|||||||||||||||||] 4.7G/15.7GB ] (lfb 143x4MB)
Swp [ ] 0.0GB/7.8GB ] (cached 0MB)
EMC [|||||||||||||] 33% 2.1GHz

GPU [|||||||||||||||||||||||||||||] 19% 1.4GHz
Dsk [#####] 13.1GB/27.4GB

[info] [Sensor] [Temp] [Power/mW] [Cur] [Avr]
UpT: 1 days 0:14:24 AO 36.00C CPU 1990 1989
FAN [|||||||||] 100% Ta=100% AUX 34.50C CV 0 0
Jetson Clocks: inactive CPU 38.00C GPU 21594 21293
NV Power[0]: MAXN GPU 42.50C SOC 3981 3929
APE: 150MHz PMIC 100.00C SYS5V 4179 4178
HW engine: Tboard 35.00C VDDRQ 1990 1970
ENC: NOT RUNNING Tdiode 39.00C
DEC: NOT RUNNING thermal 37.65C
```

Benchmarking ([link](#))^{*}

* spoilers for those following along live

Most Jetson devices have the ability to be put into different power modes

Power.mode	Time
10W	33.4 mins
15W	14.3 mins
30W 2 cores	10.8 mins
30W 4 cores	10.8 mins
30W MAX (8 cores)	7.5 mins

Note: this table represents basecalling performed when running on a portable powerbank.

Very rough comparison of Nvidia ecosystem[#]

Card	TFLOPS	Price	Power
Jetson Nano	~0.5	\$99	5-10W
Jetson Xavier NX*	6	\$399	10-15W
Jetson Xavier	11	\$699	10-30W
Geforce 1080 Ti	11.3	~\$650	215W
Geforce 2080 Ti	~13-14	~\$800-\$1000	250W
Titan RTX	16.3	~\$3000	280W
Tesla V100	16.4	~\$8500	250W

Price in USD

[#] a very small selection of what Nvidia offers obviously!

* this can fit in an existing Jetson Nano carrier boards

Community engagement leading to empowerment

... a few stories of the journey ...

So how about unlocking it for everyone?

Jean-Baka Domelevo Entfellner
@JeanBakaDE Follows you

Bioinformatics Scientist @BecAHub, Nairobi, Kenya

📍 Nairobi, Kenya 🌙 Born August 30 📅 Joined April 2019

47 Following 380 Followers

Followed by Fernanda Foertter 🇧🇷🇺🇸🇮🇳🇲🇽 and Oxford Nanopore

The unboxing and a slight snag



[View original](#)



Jean-Baka Domelevo Entfellner @JeanBakaDE
Hey Miles! Got that RTX 2070 just today! 😊
pic.twitter.com/mS1Whv2wc5

27 Aug 20, 8:59pm



[View original](#)



Jean-Baka Domelevo Entfellner @JeanBakaDE
pic.twitter.com/W8mZL0YGXc

27 Aug 20, 10:48pm

Success



[View original](#)



Jean-Baka Domelevo Entfellner @JeanBakaDE

29 Aug 20, 12:40am

I'd say it's not too bad, right? 😊 [pic.twitter.com/blqRFfqg4Z](#)

...

```
Processes: GPU GI CI PID Type Process name GPU Memory Usage
-----+----+----+----+----+----+----+----+----+
| 0 N/A N/A 1216 G /usr/lib/xorg/Xorg 18MB |
| 0 N/A N/A 1325 G /usr/bin/gnome-shell 49MB |
| 0 N/A N/A 1716 G /usr/lib/xorg/Xorg 140MB |
| 0 N/A N/A 1863 G /usr/bin/gnome-shell 102MB |
| 0 N/A N/A 4659 C guppy_basecaller 6427MB |
-----+----+----+----+----+----+----+----+----+
nanoseq@ILRIKE-021761:~$ nvidia-smi
Fri Aug 28 06:19:20 2020
+-----+-----+-----+-----+-----+-----+-----+-----+
| NVIDIA-SMI 450.51.06 Driver Version: 450.51.06 CUDA Version: 11.0 |
| GPU Name Persistence-M| Bus-Id Disp.A Volatile Uncorr. ECC | | | |
| GPU Temp Perf Pwr/Usage/Cap| Memory Usage GPU-Util Compute M. |
| Fan | | | | MIG M. |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 GeForce RTX 2070 On 00000000:00:08.0 On 99% Default |
| 60K 63C P2 174M / 175M 67391MB / 79731MB |
+-----+-----+-----+-----+-----+-----+-----+-----+
nanoseq@ILRIKE-021761:~$
```

[View original](#)

Jean-Baka Domelevo Entfellner @JeanBakaDE [pic.twitter.com/SmZS9m6t1N](#)

29 Aug 20, 6:08pm

```
nanoseq@ILRIKE-021761:~$ guppy_basecaller -l /media/nanoseq/SeqData_LTS/Data/Lablab_1470_bases_dan-dcm-cpp_hac.cfg -x auto
ONT Guppy basecalling software version 4.0.15+569a0742, client-server API version 2.1.0
config file: /opt/ont/guppy/data/dna_f9.4.1_450bps_modbases_dan-dcm-cpp_hac.cfg
model file: /opt/ont/guppy/data/featuredbase_f9.4.1_450bps_modbases_dan-dcm-cpp_hac
input path: /media/nanoseq/SeqData_LTS/Data/Lablab_1470_Run/Lablab_1470_2820962
save path: /home/nanoseq/seq_data_lablab_1470_run/Lablab_1470_2820962
chunk size: 2000
chunks per runner: 512
records per file: 4000
num basecallers: 4
gpu device: auto
kernel path:
runners per device: 4
Found 16 fast5 files to process.
Init time: 1444 ms
0% 10 20 30 40 50 60 70 80 90 100%
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Caller time: 937952 ms, Samples called: 6798788953, samples/s: 7.24846e+06
Finishing up any open output files.
Basecalling completed successfully.
nanoseq@ILRIKE-021761:~$
```

[View original](#)

Jean-Baka Domelevo Entfellner @JeanBakaDE [pic.twitter.com/kEJ501001N](#)

29 Aug 20, 6:08pm

A global cast of many

- [*Kenya*] Jean-Bakka Domelevo Entfellner ([@JeanBakaDE](#))
- [*USA*] John Erb-Downward ([@Bombanat](#))
- [*Italy*] Luigi Faino ([@lfaino](#))
- [*Switzerland*] Jürgen Hench
- [*USA*] Nvidia Clara Team ([Clara AGX site](#))

You can follow the story here: [GitHub Gist - Xavier basecalling notes](#)

The future...



POWERED BY
NVIDIA® TEGRA®

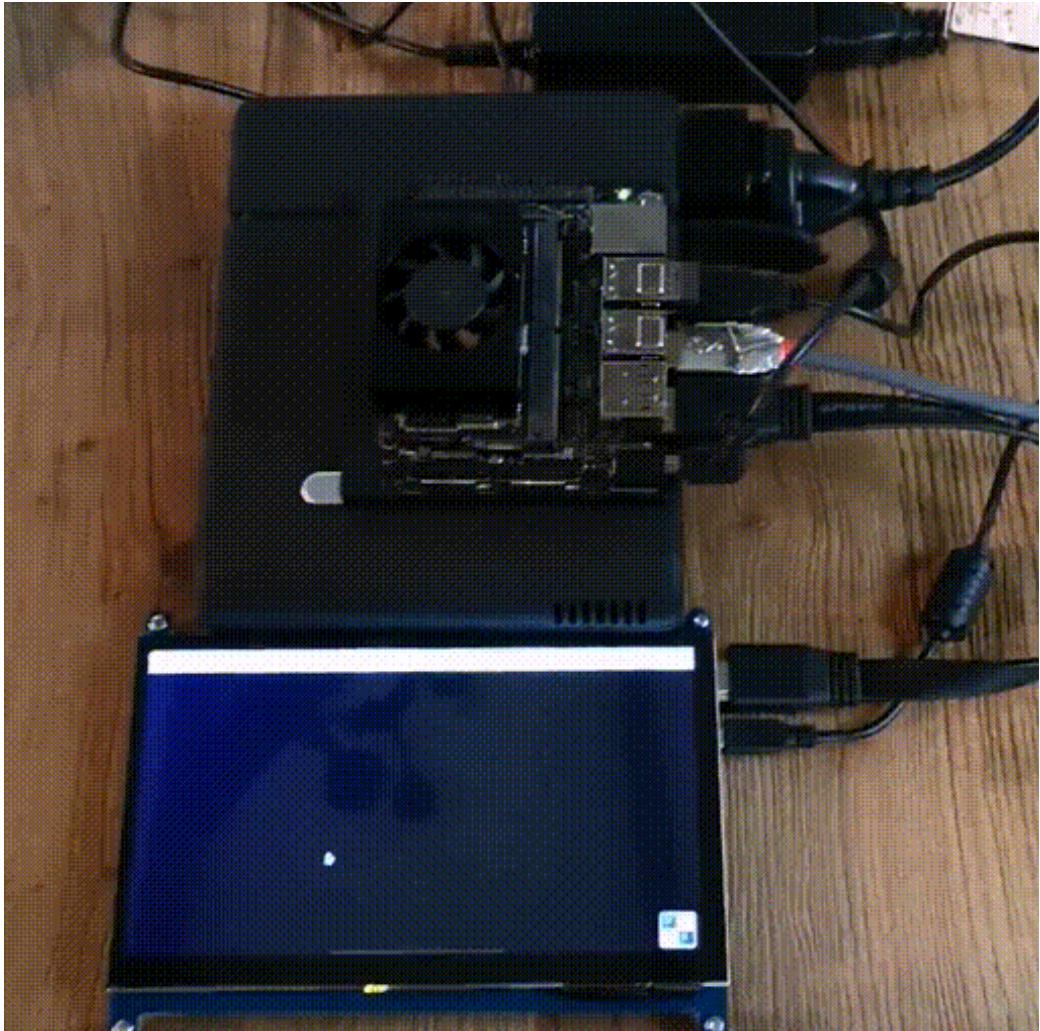
The (not so distant) future...



POWERED BY
NVIDIA® TEGRA®

E/S/R
Science for Communities

The present!



HE PŪTAIAO, HE TANGATA

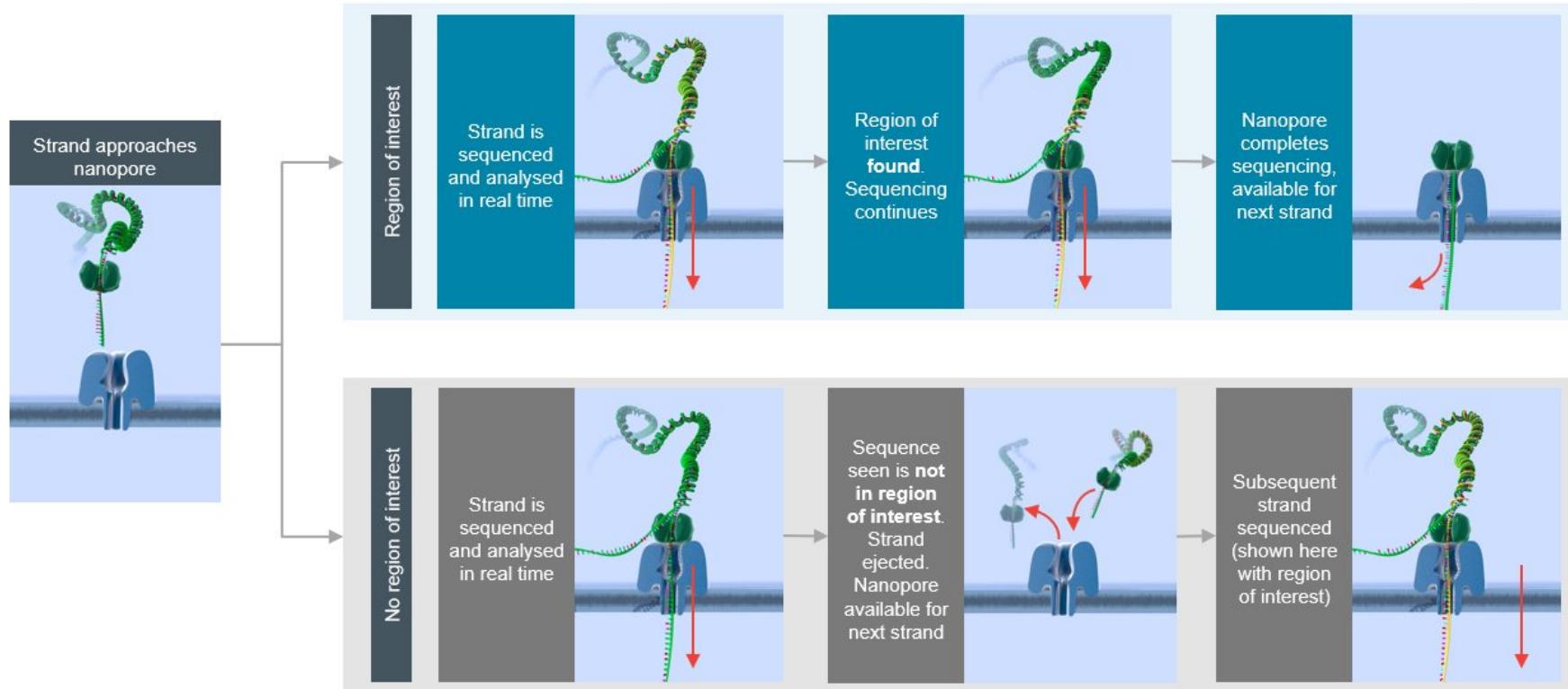
Māori science and research at ESR



E/S/R

He Pūtaiao, He Tangata

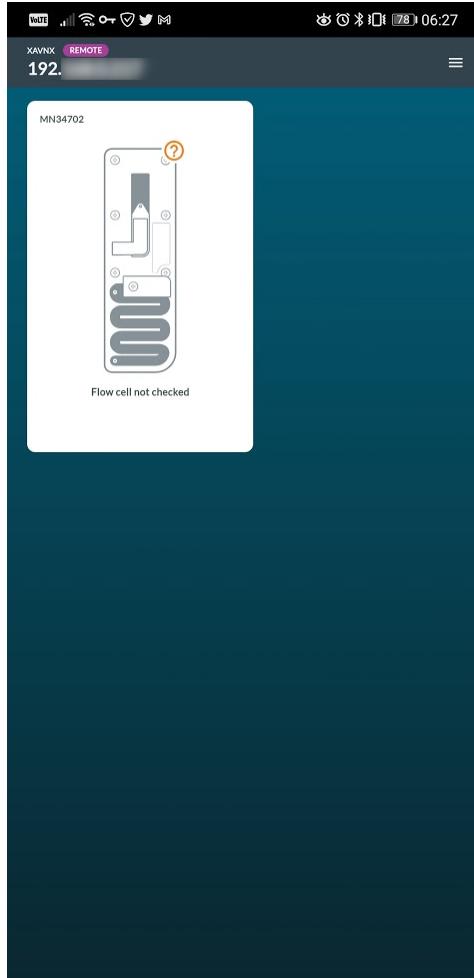
Hot off the press!



© Oxford Nanopore Technologies

Jürgen Hench, Basel, Switzerland (cancer methylation sequencing)

Hot off the press!



Can now control all of this from our pockets!

Available now on [Android](#) and [iOS](#)

Let's take a tour...

... live video content ...

Our experiences with the Jetson Xaviers

Article in the NZ Herald

- The Genius of Genomics ([link](#))
- Story also on ESR website ([link](#))

Jetson Nanopore Sequencing

- A place to collate notes and resources of our journey into porting nanopore sequencing over to accessible, portable technology
 - up to date information, including parts list and set up guide

eResearch 2020 presentation

- Assessing the potential of ~~autonomous AI devices~~ cool tech for portable real-time sequencing
 - live base-calling and presentation run from the Xavier

Xavier AGX setup and notes

- Xavier AGX unboxing and setup

Xavier base-calling benchmarking and notes

- Xavier benchmarking

If you're interested in GPUs and genomics...

Miles Benton
@miles_benton

I wrote a thing: hackmd.io/@Miles/B1U-cOM...

If you are interested in @nvidia #GPUs & compute for @nanopore MinION sequencing I put a few of my thoughts down. I'm not saying they're good thoughts, but there might be some useful info in there for people wanting to get started.

 GPU musings (with an eye on genomics) - HackMD
GPU musings (with an eye on genomics) [
hackmd.io

10:10 PM · Jan 21, 2021 from Upper Hutt City, New Zealand · Twitter for Android

[View Tweet activity](#)

20 Retweets 2 Quote Tweets 44 Likes

Reply Share Like Retweet

[GPU musings \(with an eye on genomics\)](#)

Some parting thoughts

acknowledgements (people to 'nerd' out with!)



Matt Storey
(ESR)



Joep de Ligt
(ESR)



Donia Macartney-Coxson
(ESR)

Georgia Bell (ESR)
Una Ren (ESR)
Richard Dean (ESR)

Matthew Walton (ESR)
Shane Sturrock (ESR)
Russell Smithies (ESR)

Tai Hua Nui, Tai Hua Roa

Our Impact, Our Value

