



# Another talk about CAW?

NGI Stockholm CAWterly meeting

---

Maxime Garcia

@gau

@MaxUlysse

<https://maxulysse.github.io/>

2017-10-06

SciLifeLab NGI / BarnTumörBanken

SciLifeLab

NATIONAL ATCA GENOMICS INFRASTRUCTURE

NBIS  
NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN

Karolinska Institutet

KTH  
KONSTALLEG  
ROYAL INSTITUTE OF TECHNOLOGY

Stockholms  
universitet

UPPSALA  
UNIVERSITET



# What is CAW?

A screenshot of a Google search results page for the query "CAW". The search bar at the top contains "CAW". Below the search bar, the Google logo is visible, followed by the word "CAW" and a magnifying glass icon. A horizontal navigation bar includes "All" (which is underlined), "Images", "Videos", "News", "Maps", and "More", along with "Settings" and "Tools". A message indicates "About 64 400 000 results (0,55 seconds)".

**Dictionary**

**caw**  
/kɔ:/ ⓘ

*noun*

1. the harsh cry of a rook, crow, or similar bird.

*verb*

1. utter a caw.  
"rooks cawed in the dark trees"

Translations, word origin, and more definitions

Feedback

# What is CAW?

---



<http://opensource.scilifelab.se/projects/caw/>

# What is CAW?

---



<http://opensource.scilifelab.se/projects/caw/>

- Pipeline developed at NGI



# What is CAW?

---



<http://opensource.scilifelab.se/projects/caw/>

- Pipeline developed at NGI
- In collaboration with NBIS



# What is CAW?



<http://opensource.scilifelab.se/projects/caw/>

- Pipeline developed at NGI
- In collaboration with NBIS
- Support of The Swedish Pediatric Tumor Biobank



# What does CAW do?

---



<http://opensource.scilifelab.se/projects/caw/>

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
- MuTect1, MuTect2, Strelka, and GATK HaplotypeCaller

# What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
  
- SNPs, SNVs and indels

# What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
- SNPs, SNVs and indels
- Manta

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
- SNPs, SNVs and indels
- Structural variants

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
- SNPs, SNVs and indels
- Structural variants
- ASCAT

## What does CAW do?



<http://opensource.scilifelab.se/projects/caw/>

- Tumor/Normal pair WGS analysis
- Based on GATK best practices for processing FASTQ files
- SNPs, SNVs and indels
- Structural variants
- Heterogeneity, ploidy and CNVs

# Using Singularity



<http://singularity.lbl.gov/>

<https://singularity-hub.org/>

- Docker-like containers technology
- Specific for HPC environment

# Using Singularity



<http://singularity.lbl.gov/>

<https://singularity-hub.org/>

- Docker-like containers technology
- Specific for HPC environment
- Supported by Nextflow

# Using Singularity



<http://singularity.lbl.gov/>

<https://singularity-hub.org/>

- Docker-like containers technology
- Specific for HPC environment
- Supported by Nextflow
- Better reproducibility
- Better shareability

# Using Singularity



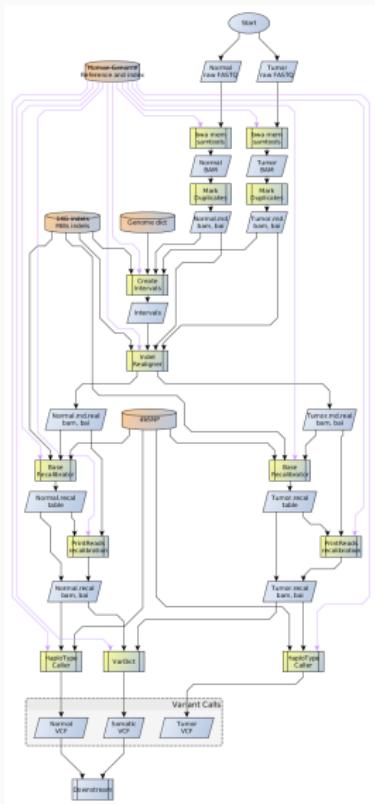
<http://singularity.lbl.gov/>

<https://singularity-hub.org/>

- Docker-like containers technology
- Specific for HPC environment
- Supported by Nextflow
- Better reproducibility
- Better shareability

Containerization is not only portable but also reproducible.

# Can CAW do more?



## Can CAW do more?

---

- Start from FASTQ files
- More or less processed BAM files
- Annotations with snpEff and/or VEP
- Reports with MultiQC

## Can CAW do more?

---

- Start from FASTQ files
- More or less processed BAM files
- Annotations with snpEff and/or VEP
- Reports with MultiQC
- Choose between GRCh37 or GRCh38

## Can CAW do more?

---

- Start from FASTQ files
- More or less processed BAM files
- Annotations with snpEff and/or VEP
- Reports with MultiQC
- Choose between GRCh37 or GRCh38
- We can also process normal only samples

Can CAW do more?

OH, REALLY ?



PLEASE, TELL ME MORE

# CAW to replace Piper into production

---



- Kick off meeting was 2017-10-04

# CAW to replace Piper into production

---



- Kick off meeting was 2017-10-04
- Normal only processing of samples

# CAW to replace Piper into production

---



- Kick off meeting was 2017-10-04
- Normal only processing of samples
- Handles GRCh38 and GRCh37

# CAW to replace Piper into production

---



- Kick off meeting was 2017-10-04
- Normal only processing of samples
- Handles GRCh38 and GRCh37
- Faster than Piper

# CAW to replace Piper into production

---



- Kick off meeting was 2017-10-04
- Normal only processing of samples
- Handles GRCh38 and GRCh37
- Faster than Piper
- Has already been used to process the 1000 samples from SweGen

## The List of People Involved

---

Sebastian DiLorenzo	Monica Nistèr
Jesper Eisfeldt	Björn Nystedt
Maxime Garcia	Pall Olason
Szilveszter Juhos	Markus Ringnér
Max Käller	Pelin Sahlén
Malin Larsson	Johanna Sandgren
Marcel Martin	Teresita Díaz De Ståhl
Markus Mayrhofer	

## Where to find us?

- We are on the SciLifeLab Slack   
#cancer-pipeline

# Where to find us?

- We are on the SciLifeLab Slack   
`#cancer-pipeline`
- We have a gitter channel  
<https://gitter.im/SciLifeLab/CAW>

# Where to find us?

- We are on the SciLifeLab Slack   
`#cancer-pipeline`
- We have a gitter channel  
<https://gitter.im/SciLifeLab/CAW>
- Our code is hosted on Github   
<https://github.com/SciLifeLab/CAW>

# Any questions?

