

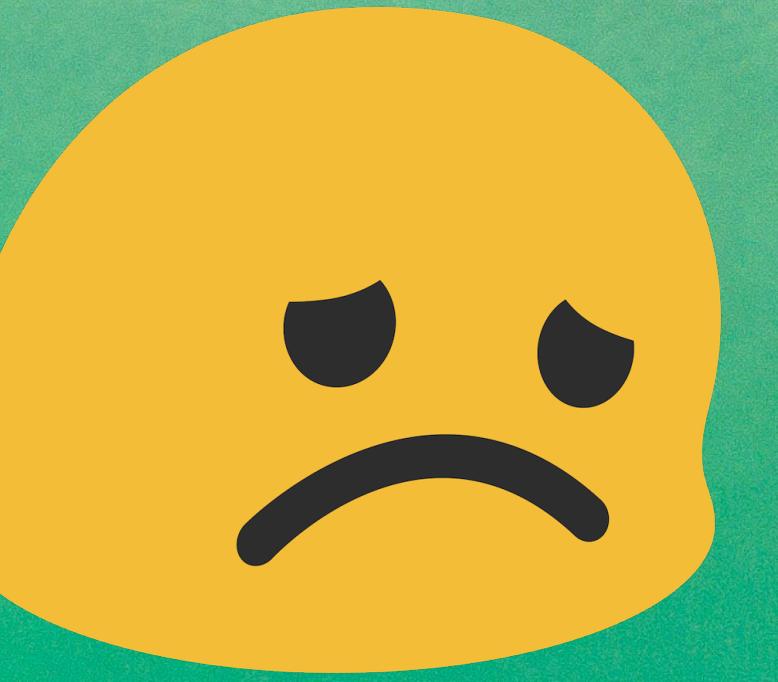
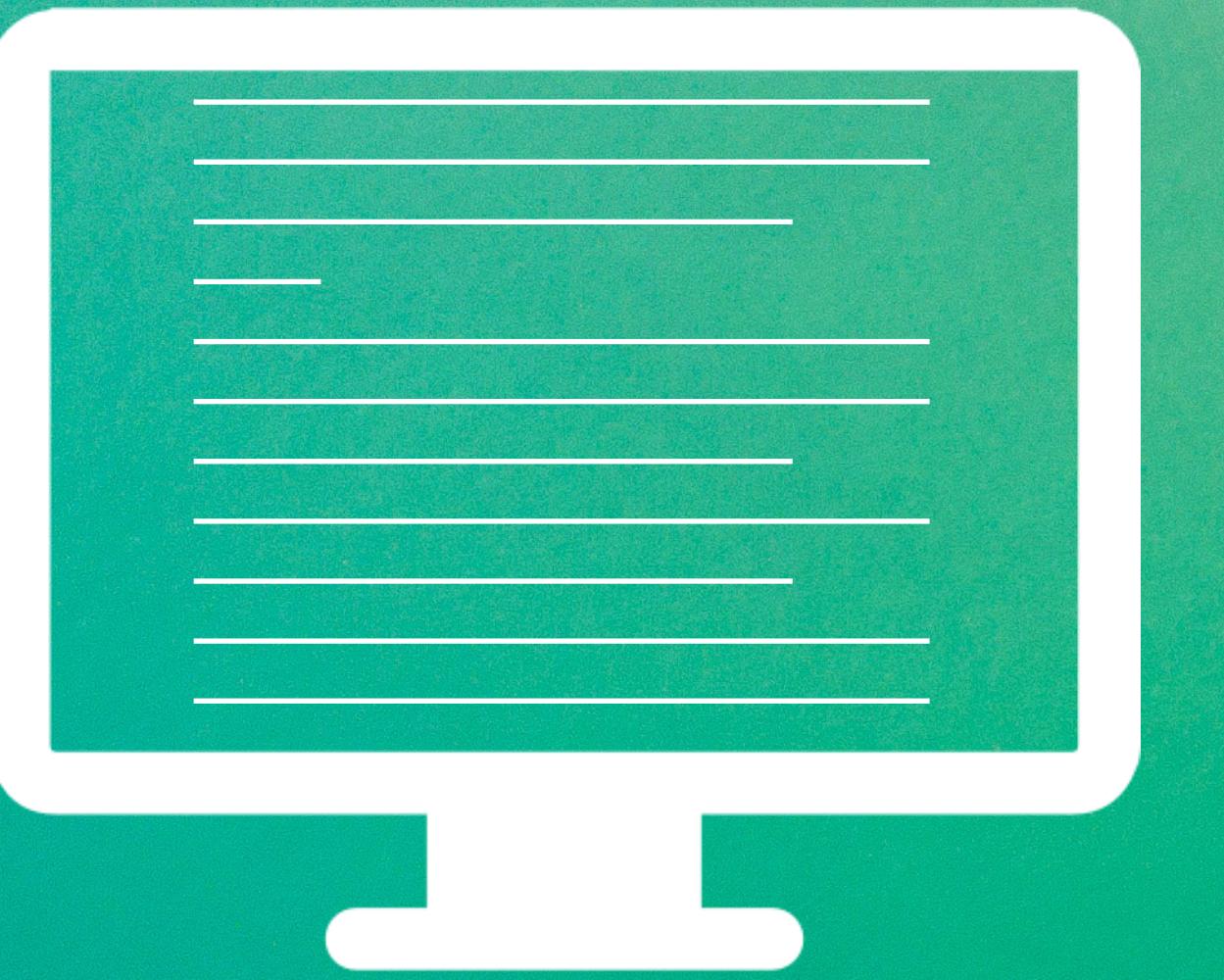
# **nf-core/ bytesize**

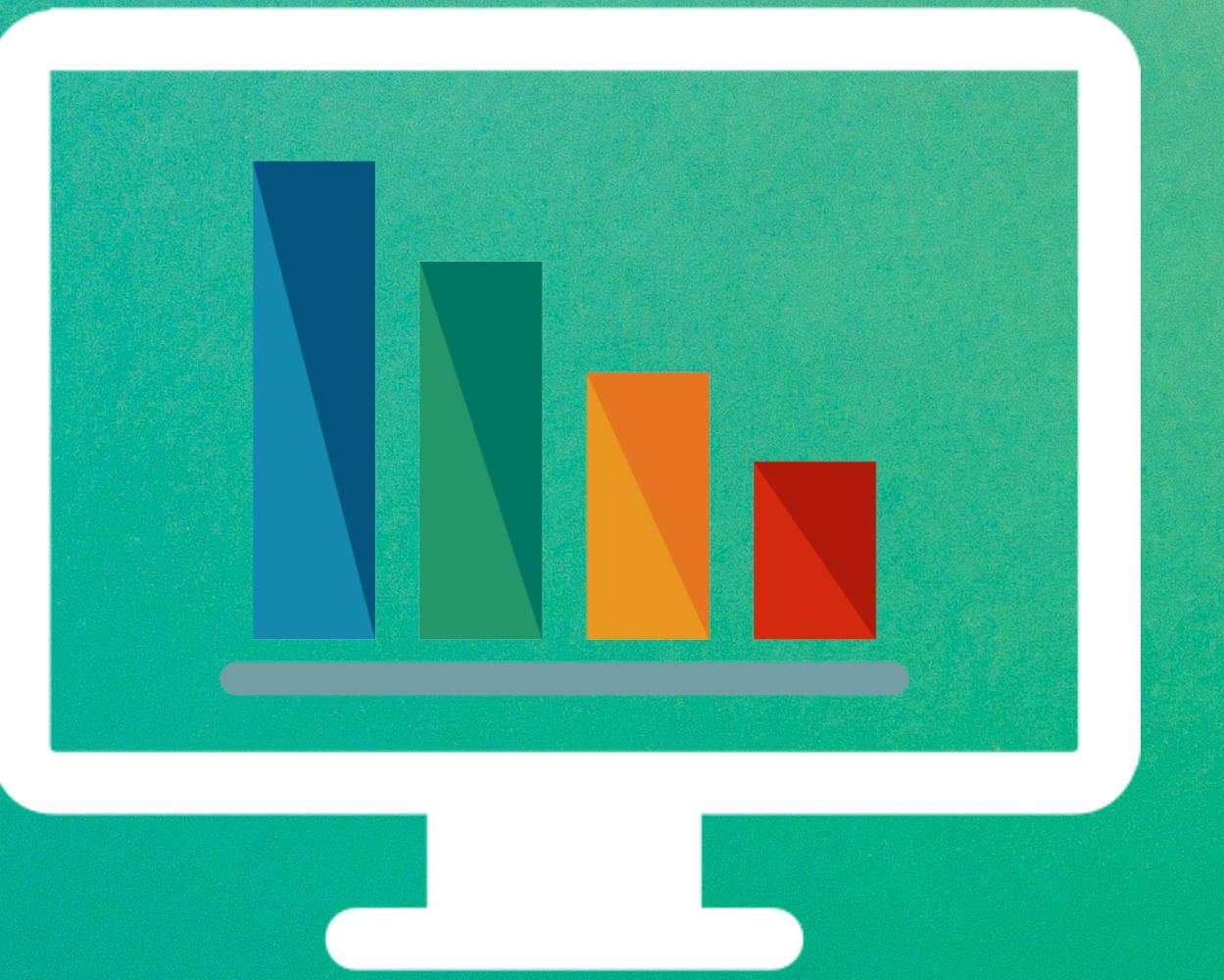


#36

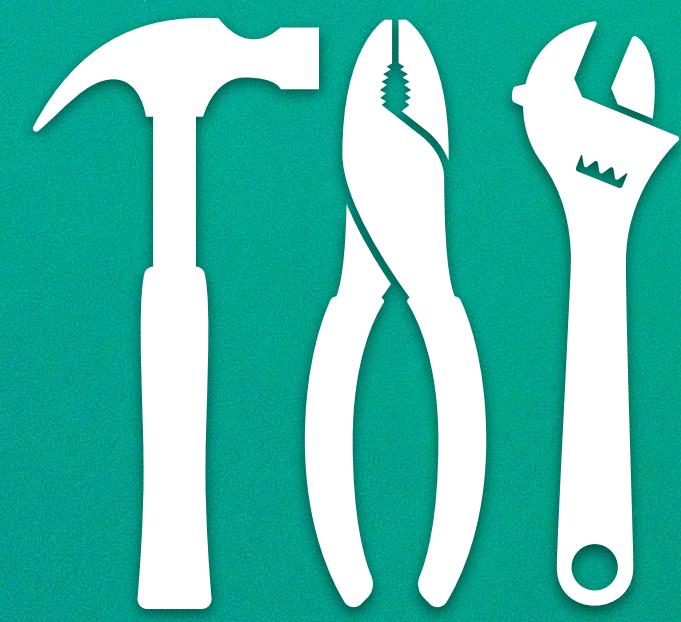
## Customising MultiQC

Phil Ewels, SciLifeLab Sweden





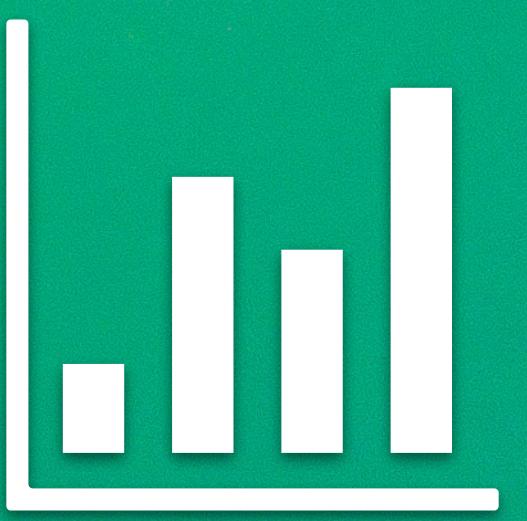
# MultiQC



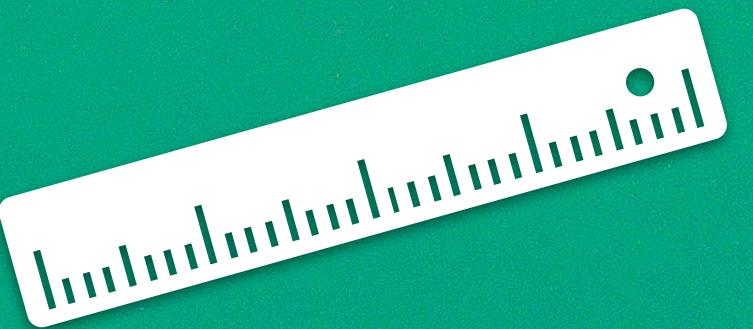
Multiple  
tools



Multiple  
samples



One  
report



Standardised  
output

# MultiQC

- 1 Install MultiQC
- 2 Run MultiQC
- 3 Read the report

```
pip install multiqc
```

```
multiqc .
```



# MultiQC



114

Tools supported

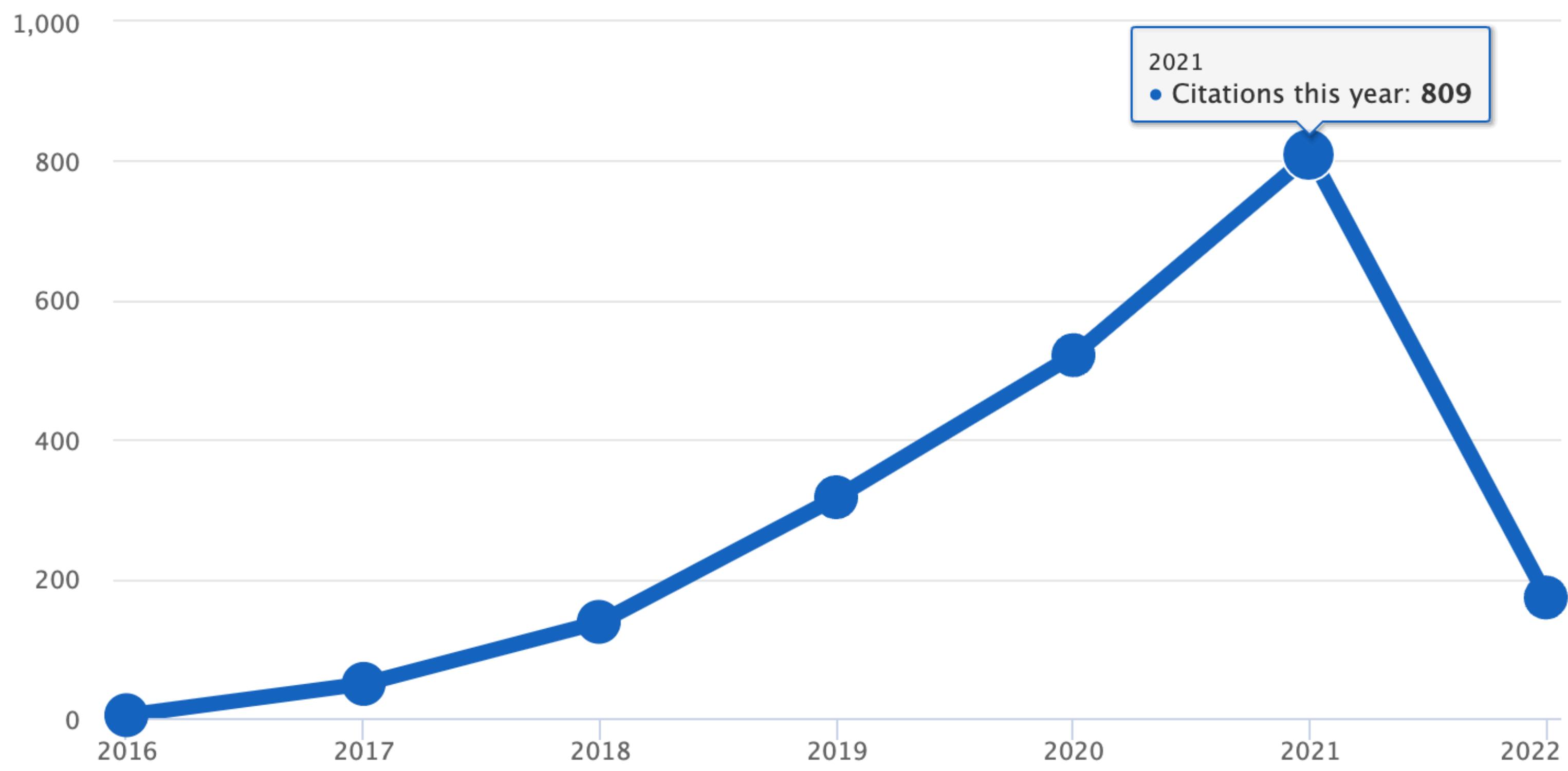
2,014

Citations





### Citations by year



# MultiQC



1,092

Issues  
(148 open)

551

Pull requests  
(35 open)



# MultiQC



0.45

Issues per day

0.23

PRs per day



The background of the image is a sky at dusk or dawn, featuring a gradient from deep blue at the top to a warm orange and yellow near the horizon. Scattered across the sky are wispy, pinkish-orange clouds.

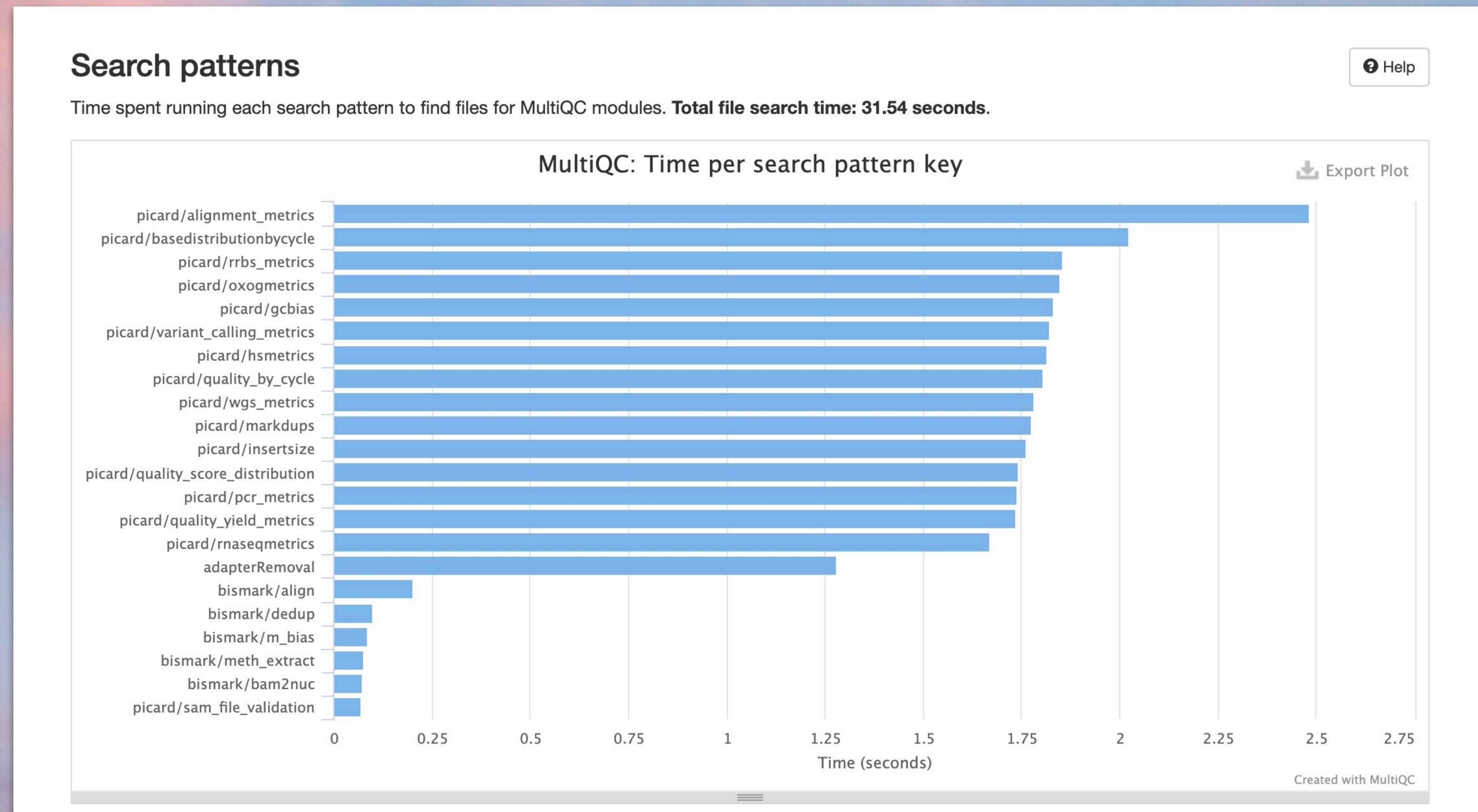
For pipeline developers

# Run time optimisation

```
multiqc --profile-runtime .
```

```
[INFO] multiqc : MultiQC complete
[INFO] multiqc : Run took 35.28 seconds
[INFO] multiqc : - 31.01s: Searching files
[INFO] multiqc : - 1.75s: Running modules
[INFO] multiqc : - 0.96s: Compressing report data
[INFO] multiqc : For more information, see the 'Run Time' section
```

# Run time optimisation



# Run time optimisation

- Only run the modules you need
- Optimise file search patterns:
  - Only run the sub-modules you need
  - Use filename patterns

<https://multiqc.info/docs/#optimising-run-time>

# Common configs

- Change the order of modules: `top_modules`
- Run modules multiple times: `module_order`
- Clean up sample names: `extra_fn_clean_exts`
- No beeswarm plots: `max_table_rows`

# More of these please

- No analysis path: `show_analysis_paths`
- Comments for modules: `section_comments`
- Report logos: `custom_logo`
- Customised plots: `custom_plot_config`
- Hide table columns: `table_columns_visible`

# Custom content

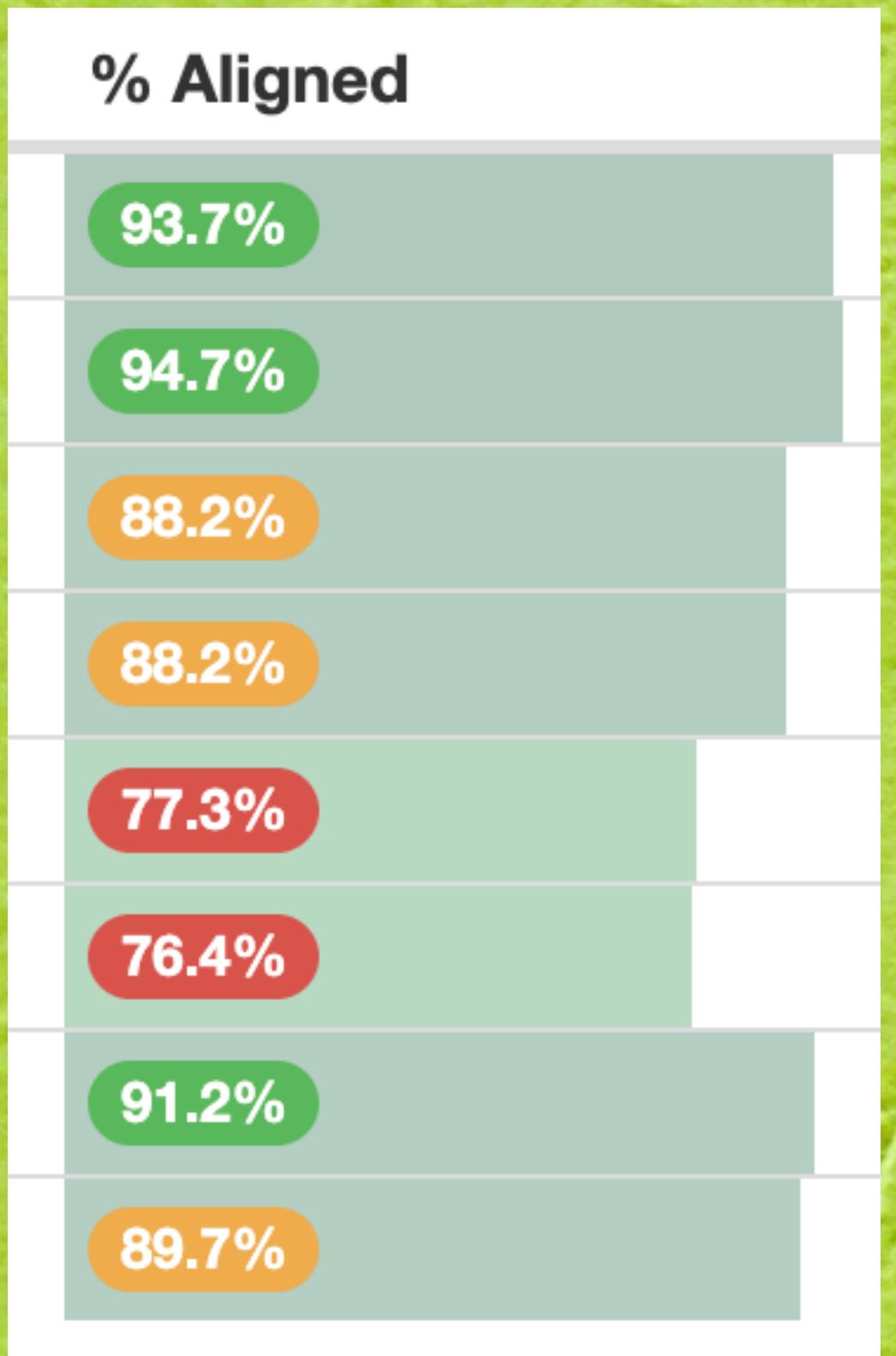
- Output from pipeline scripts
- Custom HTML
- Images (run-level only please!)
- Append to your filename: **\_mqc.json**
- Use the config to your advantage
- Examples: [https://github.com/ewels/MultiQC\\_TestData/](https://github.com/ewels/MultiQC_TestData/)

# For pipeline users

- Supply your own config: `--multiqc_config`
- Additional to the pipeline config



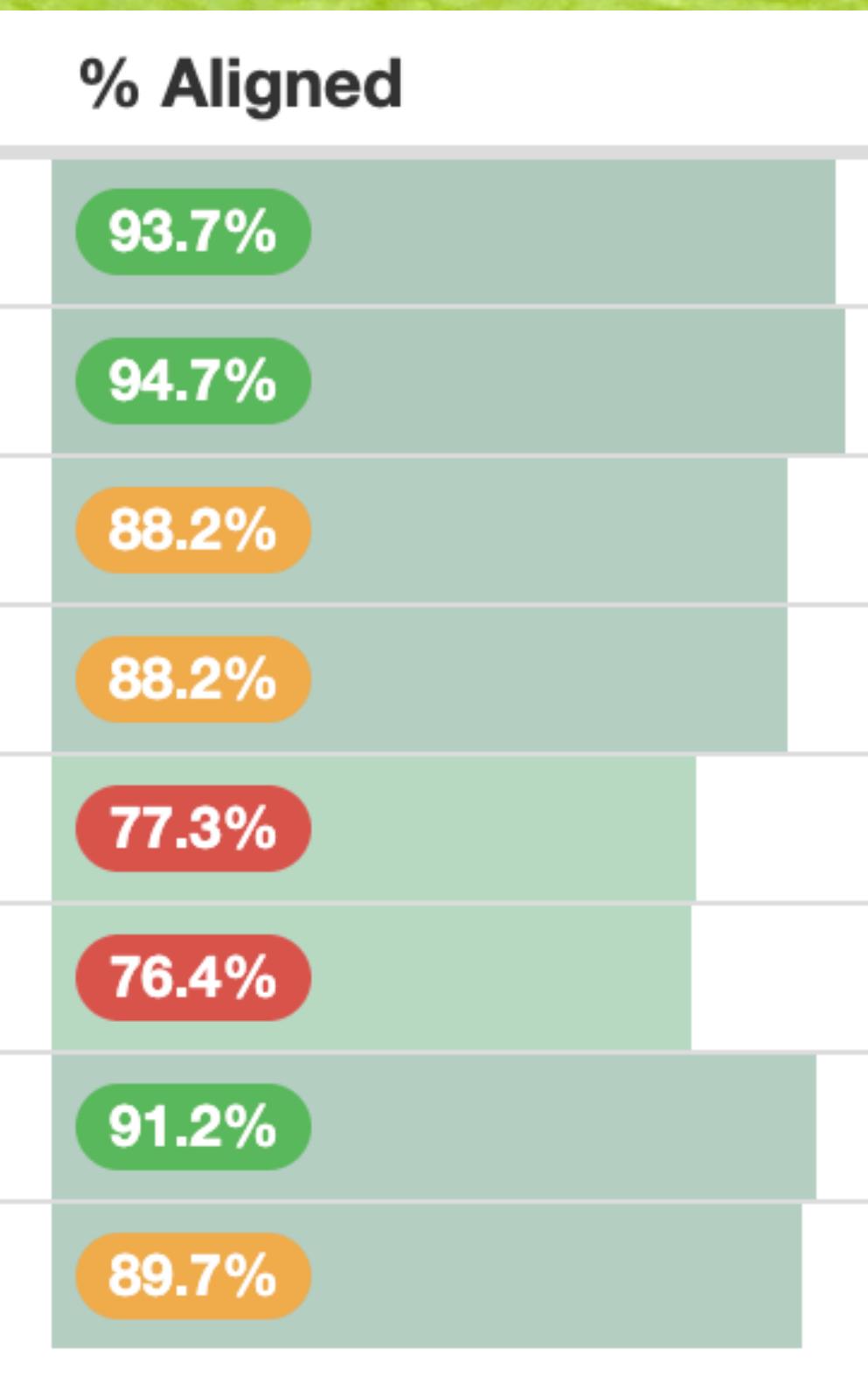
# Conditional Formatting



```
table_cond_formatting_rules:  
mqc-generalstats-uniquely_mapped_percent:  
pass:  
  - gt: 90  
warn:  
  - lt: 90  
fail:  
  - lt: 80
```



# Conditional Formatting



# Project Information

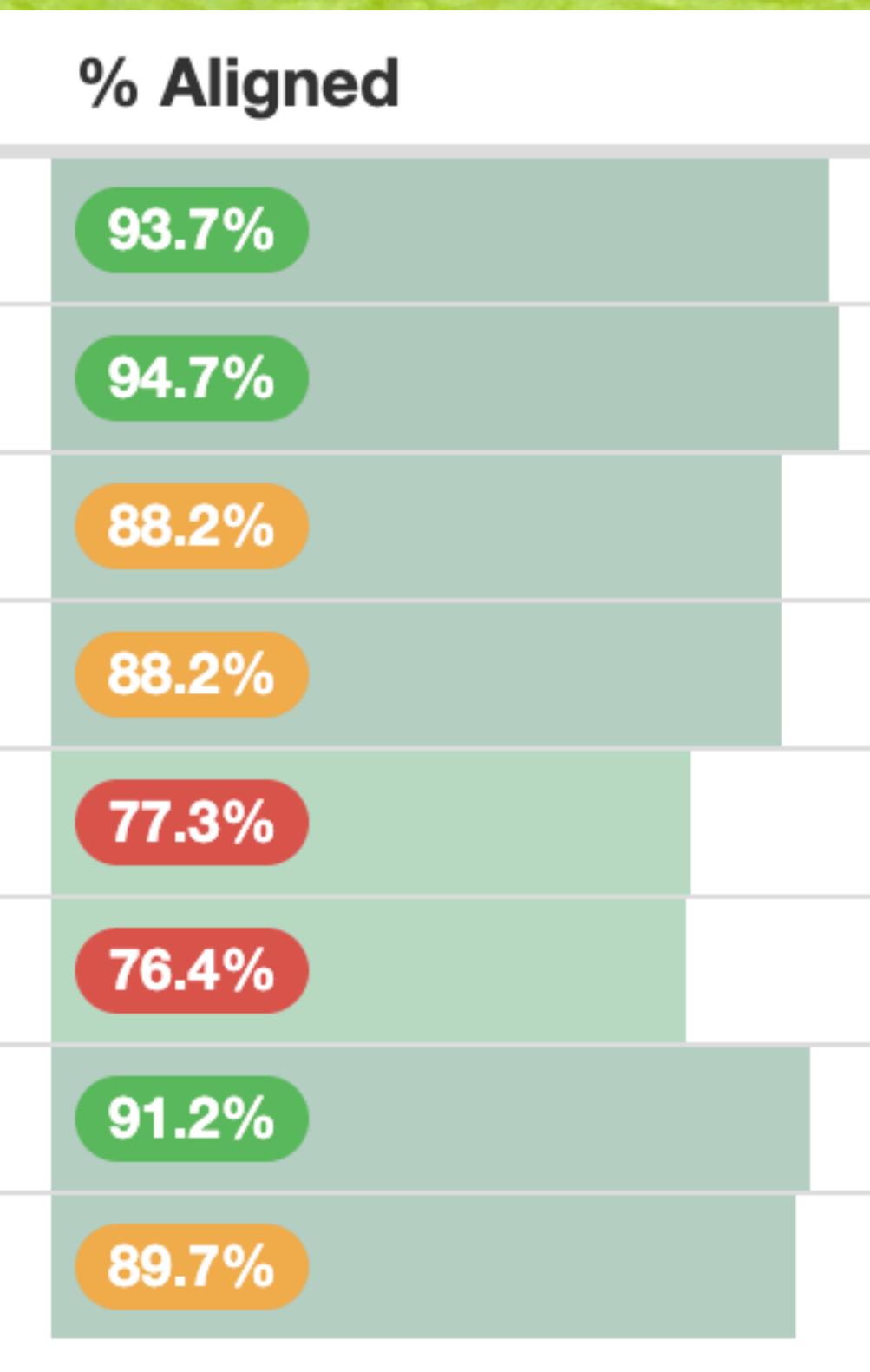
Title + subtitle

Comments

Custom sample  
names



# Conditional Formatting



# Project Information

Title + subtitle

Comments

Custom sample  
names

# Style and theme

Custom logo

Custom CSS

Custom template

M P1234: Test\_NGI\_Project: Multi +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC**  
v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

**MultiQC**

**P1234: Test\_NGI\_Project**

This is an example project. All identifying data has been removed.

This report has been generated by the [NGI-RNAseq](#) analysis pipeline. Some additional functionality is added to the report using the [MultiQC\\_NGI](#) plugin.

Contact E-mail: [phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)  
Application Type: RNA-seq  
Sequencing Platform: HiSeq 2500 High Output V4  
Sequencing Setup: 2x125  
Reference Genome: hg19

Report generated on 2022-02-08, 23:03 based on data in: /Users/phil/GitHub/MultiQC/website/public\_html/examples/ngi-rna

[NGI names](#) [User supplied names](#)

[Welcome!](#) Not sure where to start? [Watch a tutorial video](#) (6:06) [don't show again](#) X

**General Statistics**

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed	% Dups	% GC	M Seqs
P1234_1001	68.2%	22.8	10.3%	71.3%	49%	33.7
P1234_1002	67.9%	20.9	10.7%	68.8%	50%	31.1

title

M P1234: Test\_NGI\_Project: Multi +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC**  
v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

**MultiQC**

**P1234: Test NGI Project**

This is an example project. All identifying data has been removed.

This report has been generated by the [NGI-RNAseq](#) analysis pipeline. Some additional functionality is added to the report using the [MultiQC\\_NGI](#) plugin.

Contact E-mail: [phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)  
Application Type: RNA-seq  
Sequencing Platform: HiSeq 2500 High Output V4  
Sequencing Setup: 2x125  
Reference Genome: hg19

Report generated on 2022-02-08, 23:03 based on data in: /Users/phil/GitHub/MultiQC/website/public\_html/examples/ngi-rna

[NGI names](#) [User supplied names](#)

[Welcome!](#) Not sure where to start? [Watch a tutorial video](#) (6:06) [don't show again](#) X

**General Statistics**

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed	% Dups	% GC	M S
P1234_1001	68.2%	22.8	10.3%	71.3%	49%	33.7
P1234_1002	67.9%	20.9	10.7%	68.8%	50%	31.1

subtile

M P1234: Test\_NGI\_Project: Multi +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

MultiQC v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

# MultiQC

## P1234: Test\_NGI\_Project

This is an example project. All identifying data has been removed.

This report has been generated by the [NGI-RNAseq](#) analysis pipeline. Some additional functionality is added to the report using the [MultiQC\\_NGI](#) plugin.

Contact E-mail: [phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)  
Application Type: RNA-seq  
Sequencing Platform: HiSeq 2500 High Output V4  
Sequencing Setup: 2x125  
Reference Genome: hg19

Report generated on 2022-02-08, 23:03 based on data in: /Users/phil/GitHub/MultiQC/website/public\_html/examples/ngi-rna

NGI names User supplied names

Welcome! Not sure where to start? Watch a tutorial video (6:06) don't show again X

### General Statistics

Copy table Configure Columns Plot Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed
P1234_1001	68.2%	22.8	10.3%
P1234_1002	67.9%	20.9	10.7%

report\_comment

M P1234: Test\_NGI\_Project: Multi +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC**  
v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

# MultiQC

## P1234: Test\_NGI\_Project

This is an example project. All identifying data has been removed.

This report has been generated by the [NGI-RNAseq](#) analysis pipeline. Some additional functionality is added to the report using the [MultiQC\\_NGI](#) plugin.

Contact E-mail: [phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)  
Application Type: RNA-seq  
Sequencing Platform: HiSeq 2500 High Output V4  
Sequencing Setup: 2x125  
Reference Genome: hg19

Report generated on 2022-02-08, 23:03 based on data in: /Users/phil/GitHub/MultiQC/website/public\_html/examples/ngi-rna

[NGI names](#) [User supplied names](#)

[Welcome!](#) Not sure where to start? [Watch a tutorial video](#) (6:06)

### General Statistics

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed
P1234_1001	68.2%	22.8	10.3%
P1234_1002	67.9%	20.9	10.7%

custom\_logo

custom\_logo\_url

custom\_logo\_title

M P1234: Test\_NGI\_Project: Multi +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC**  
v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

# MultiQC

## P1234: Test\_NGI\_Project

This is an example project. All identifying data has been removed.

This report has been generated by the [NGI-RNAseq](#) analysis pipeline. Some additional functionality is added to the report using the [MultiQC\\_NGI](#) plugin.

Contact E-mail: [phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)  
Application Type: RNA-seq  
Sequencing Platform: HiSeq 2500 High Output V4  
Sequencing Setup: 2x125  
Reference Genome: hg19

Report generated on 2022-02-08, 23:03 based on data in: /Users/phil/GitHub/MultiQC/website/public\_html/examples/ngi-rna

[NGI names](#) [User supplied names](#)

[Welcome!](#) Not sure where to start? [Watch a tutorial video](#) (6:06) [don't show again](#) X

### General Statistics

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trim
P1234_1001	68.2%	22.8	10.3%
P1234_1002	67.9%	20.9	10.7%

report\_header\_info

M P1234: Test\_NGI\_Project: Multi X +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC** v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

**NGI names** **User supplied names**

**Welcome!** Not sure where to start? Watch a tutorial video (6:06) don't show again X

**General Statistics**

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed	% Dups	% GC	M Seqs
P1234_1001	68.2%	22.8	10.3%	71.3%	49%	33.7
P1234_1002	67.9%	20.9	10.7%	68.8%	50%	31.1
P1234_1003	64.7%	21.7	11.0%	70.9%	51%	33.7
P1234_1004	55.2%	17.0	13.2%	73.2%	52%	31.2
P1234_1005	53.0%	17.7	15.9%	75.6%	51%	33.8
P1234_1006	52.7%	16.1	14.1%	73.8%	52%	30.8
P1234_1007	33.0%	7.0	32.0%	80.5%	52%	21.8
P1234_1008	27.5%	4.3	44.2%	80.6%	50%	16.7
P1234_1009	52.3%	10.5	20.9%	64.2%	46%	20.5
P1234_1010	44.2%	4.9	24.9%	80.7%	48%	11.4
P1234_1011	42.2%	11.6	27.1%	44.2%	49%	28.0
P1234_1012	40.6%	10.6	27.1%	45.2%	51%	26.6
P1234_1013	60.8%	35.6	14.0%	51.6%	46%	59.0
P1234_1014	64.7%	30.1	12.2%	49.2%	46%	46.9
P1234_1015	48.0%	6.3	23.5%			

Toolbox

A

D

?

**-- sample-names**

M P1234: Test\_NGI\_Project: Multi X +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

**MultiQC** v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

**NGI names** **User supplied names**

**Welcome!** Not sure where to start? Watch a tutorial video (6:06) don't show again X

**General Statistics**

Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed	% Dups	% GC	M Seqs
Sample 1, Rep 1	68.2%	22.8	10.3%	71.3%	49%	33.7
Sample 1, Rep 2	67.9%	20.9	10.7%	68.8%	50%	31.1
Sample 1, Rep 3	64.7%	21.7	11.0%	70.9%	51%	33.7
Sample 2, Rep 1	55.2%	17.0	13.2%	73.2%	52%	31.2
Sample 2, Rep 2	53.0%	17.7	15.9%	75.6%	51%	33.8
Sample 2, Rep 3	52.7%	16.1	14.1%	73.8%	52%	30.8
Sample 3, Rep 1	33.0%	7.0	32.0%	80.5%	52%	21.8
Sample 3, Rep 2	27.5%	4.3	44.2%	80.6%	50%	16.7
Sample 4, Rep 1	52.3%	10.5	20.9%	64.2%	46%	20.5
Sample 4, Rep 2	44.2%	4.9	24.9%	80.7%	48%	11.4
Sample 5, Rep 1	42.2%	11.6	27.1%	44.2%	49%	28.0
Sample 5, Rep 2	40.6%	10.6	27.1%	45.2%	51%	26.6
Sample 6, Rep 1	60.8%	35.6	14.0%	51.6%	46%	59.0
Sample 6, Rep 2	64.7%	30.1	12.2%	49.2%	46%	46.9
Sample 7, Rep 1	48.0%	6.3	23.5%			

Toolbox

A

?

-- sample-names

M P1234: Test\_NGI\_Project: Multi X +

multiqc.info/examples/ngi-rna/Test\_NGI\_Project\_multiqc\_report.html#

MultiQC v1.12

P1234: Test\_NGI\_Project

General Stats

edgeR: Sample Similarity

MDS Plot

STAR

Cutadapt

Filtered Reads

Trimmed Sequence Lengths (3')

FastQC

Sequence Counts

Sequence Quality Histograms

Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Overrepresented sequences

Adapter Content

Status Checks

NGI-RNAseq Software Versions

NGI names User supplied names

Welcome! Not sure where to start? Watch a tutorial video (6:06)

General Statistics

Copy table Configure Columns Plot Showing 22/22 rows and 6/8 columns.

Sample Name	% Aligned	M Aligned	% BP Trimmed	% Dups
Sample 1, Rep 1	68.2%	22.8	10.3%	71.3%
Sample 1, Rep 2	67.9%	20.9	10.7%	68.8%
Sample 1, Rep 3	64.7%	21.7	11.0%	70.9%
Sample 2, Rep 1	55.2%	17.0	13.2%	73.2%
Sample 2, Rep 2	53.0%	17.7	15.9%	75.6%
Sample 2, Rep 3	52.7%	16.1	14.1%	73.8%
Sample 3, Rep 1	33.0%	7.0	32.0%	80.5%
Sample 3, Rep 2	27.5%	4.3	44.2%	80.6%
Sample 4, Rep 1	52.3%	10.5	20.9%	64.2%
Sample 4, Rep 2	44.2%	4.9	24.9%	80.7%
Sample 5, Rep 1	42.2%	11.6	27.1%	44.2%
Sample 5, Rep 2	40.6%	10.6	27.1%	45.2%
Sample 6, Rep 1	60.8%	35.6	14.0%	51.6%
Sample 6, Rep 2	64.7%	30.1	12.2%	49.2%
Sample 7, Rep 1	48.0%	6.3	23.5%	

MultiQC Toolbox

Rename Samples Apply

From To +

Click here for bulk input.

Regex mode off help Clear

A

P1234\_1020 > Sample 9, Rep 1

P1234\_1022 > Failed Library

P1234\_1019 > Sample 9, Rep 1

P1234\_1018 > Sample 8, Rep 1

P1234\_1013 > Sample 6, Rep 1

P1234\_1012 > Sample 5, Rep 1

P1234\_1011 > Sample 5, Rep 1

P1234\_1010 > Sample 4, Rep 1

P1234\_1017 > Sample 8, Rep 1

P1234\_1016 > Sample 7, Rep 1

P1234\_1015 > Sample 7, Rep 1

P1234\_1014 > Sample 6, Rep 1

P1234\_1021 > Sample 10, Rep 1

P1234\_1004 > Sample 2, Rep 1

P1234\_1005 > Sample 2, Rep 1

P1234\_1006 > Sample 2, Rep 1

P1234\_1007 > Sample 3, Rep 1

-- sample-names

M MultiQC Report

file:///Users/philewels/Desktop/MultiQC\_testing/star/multiqc\_report/multiqc\_report.html#top

Toggle navigation

v0.9.dev0

General Stats

NEW featureCounts

STAR

Cutadapt

FastQ Screen

FastQC

Sequence Quality Histograms

NEW Per Sequence Quality Scores

Per Base Sequence Content

Per Sequence GC Content

Per Base N Content

Sequence Length Distribution

Sequence Duplication Levels

Adapter Content

MultiQC

A modular tool to aggregate results from bioinformatics tools

Report generated on 2016-10-17, 15:10 based on data in: /Users/philewels/Desktop/MultiQC\_testing/star/data

General Statistics

Copy table Configure Columns Showing 6/6 rows and 7/10 columns.

Sample Name	% Assigned	M Assigned	% Aligned	M Aligned	% Trimmed	% Dups	% GC	Length	% Failed	M S
SRR1067503_1	2.4%	0.9	61.2%	19.3	2.1%	2.9%	46%	35	25%	30.5
SRR1067505_1	7.4%	1.5	79.1%	14.2	3.5%	7.8%	47%	35	25%	18.0
SRR1067510_1	1.1%	0.6	50.6%	17.4	2.0%	1.4%	40%	35	25%	34.3
SRR1067514_1	5.7%	1.9	70.2%	23.6	3.1%	6.6%	46%	35	17%	33.6
SRR1067519_1	3.2%	0.9	81.1%	19.9	2.3%	5.8%	42%	35	25%	24.6
SRR1067522_1	1.4%	0.7	61.8%	22.0	1.5%					

featureCounts

Toolbox

--template geo

# Looking to the future

```
import multiqc  
multiqc.run('~/data')
```



Looking to the future





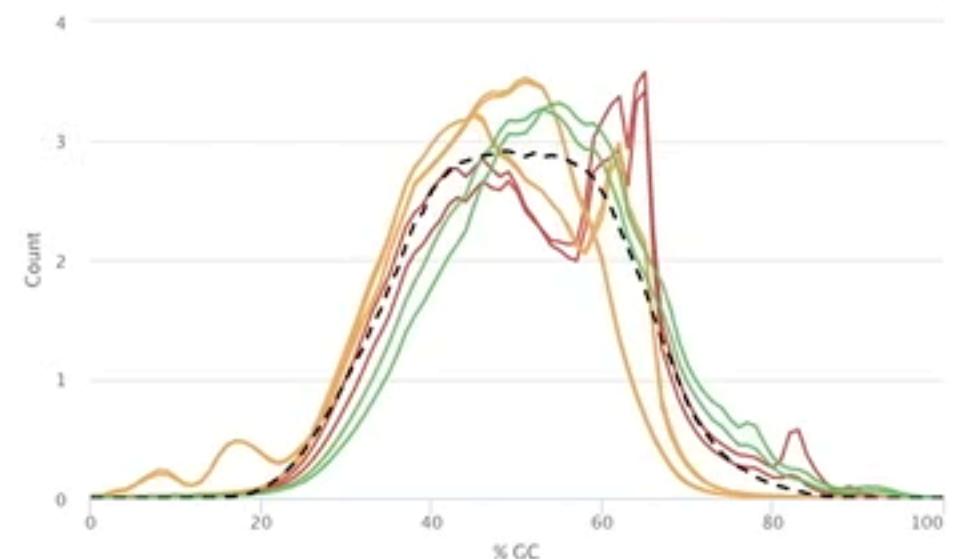
MegaQC is a web application that collects results from multiple runs of MultiQC and allows bulk visualisation.

See the MultiQC website for installation instructions and documentation: <http://multiqc.info/megaqc/>

Data available from  
38 samples, 3 reports

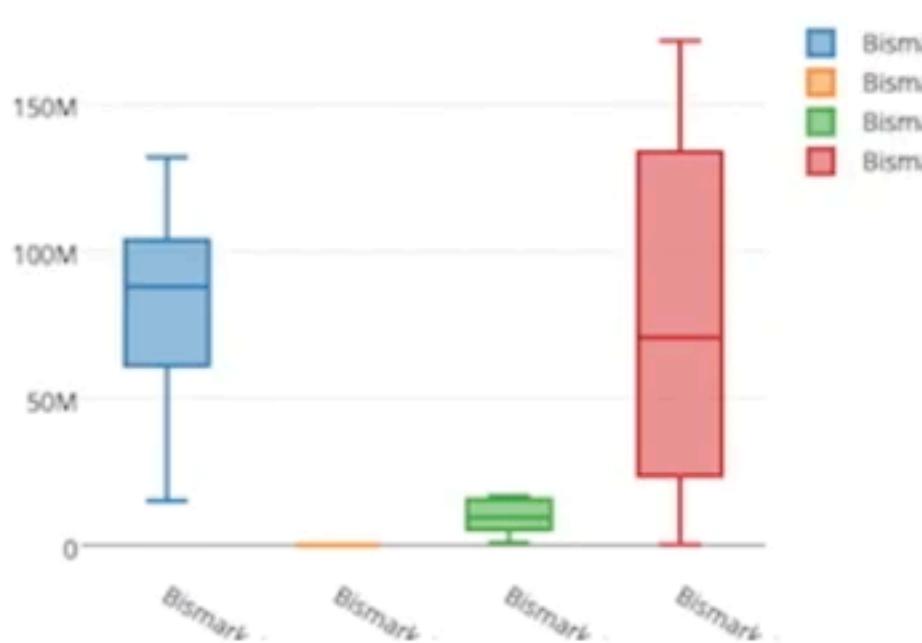
## Choose a plot type

### Report Plot



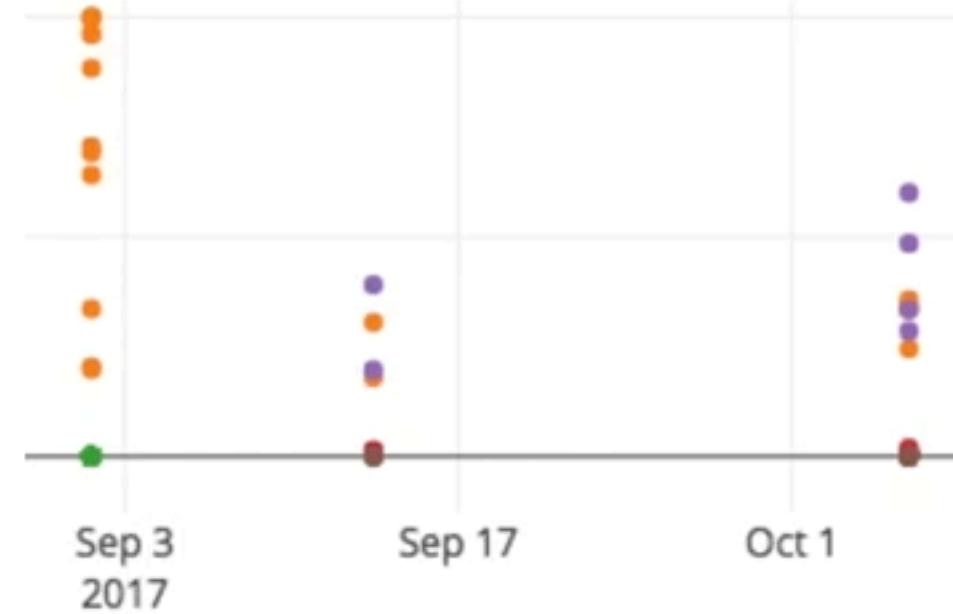
Replicate a plot from a submitted MultiQC report, using all samples in the database.

### Data Distributions



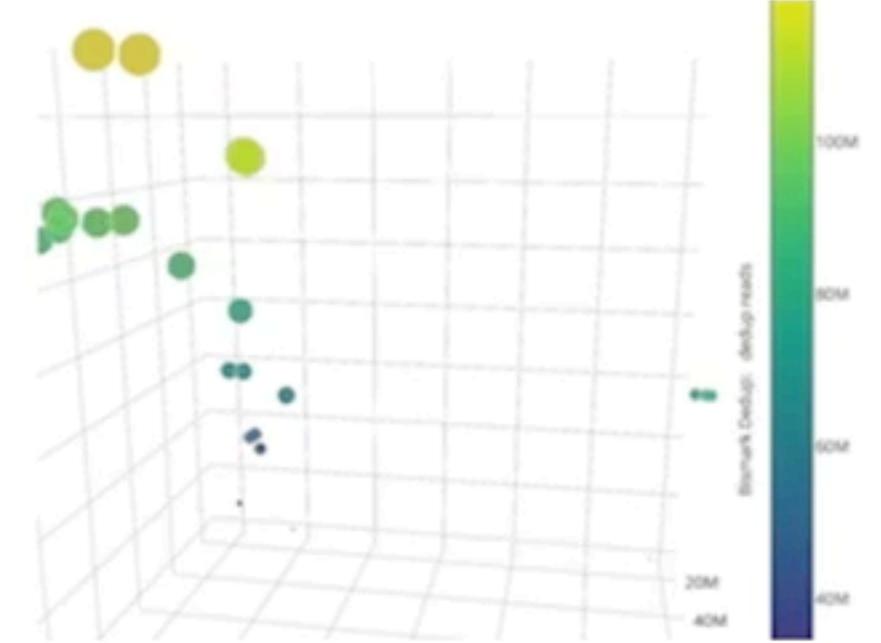
View histograms and box plots of values from MultiQC reports as a histogram.

### Data Trend



See how values from MultiQC reports change over time to find trends.

### Compare Data



Compare values from MultiQC reports against each other to look for correlations.



 <http://multiqc.info>



 <https://github.com/ewels/MultiQC>

 <https://gitter.im/ewels/MultiQC>

Phil Ewels  
[phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)

 tallphil

 ewels