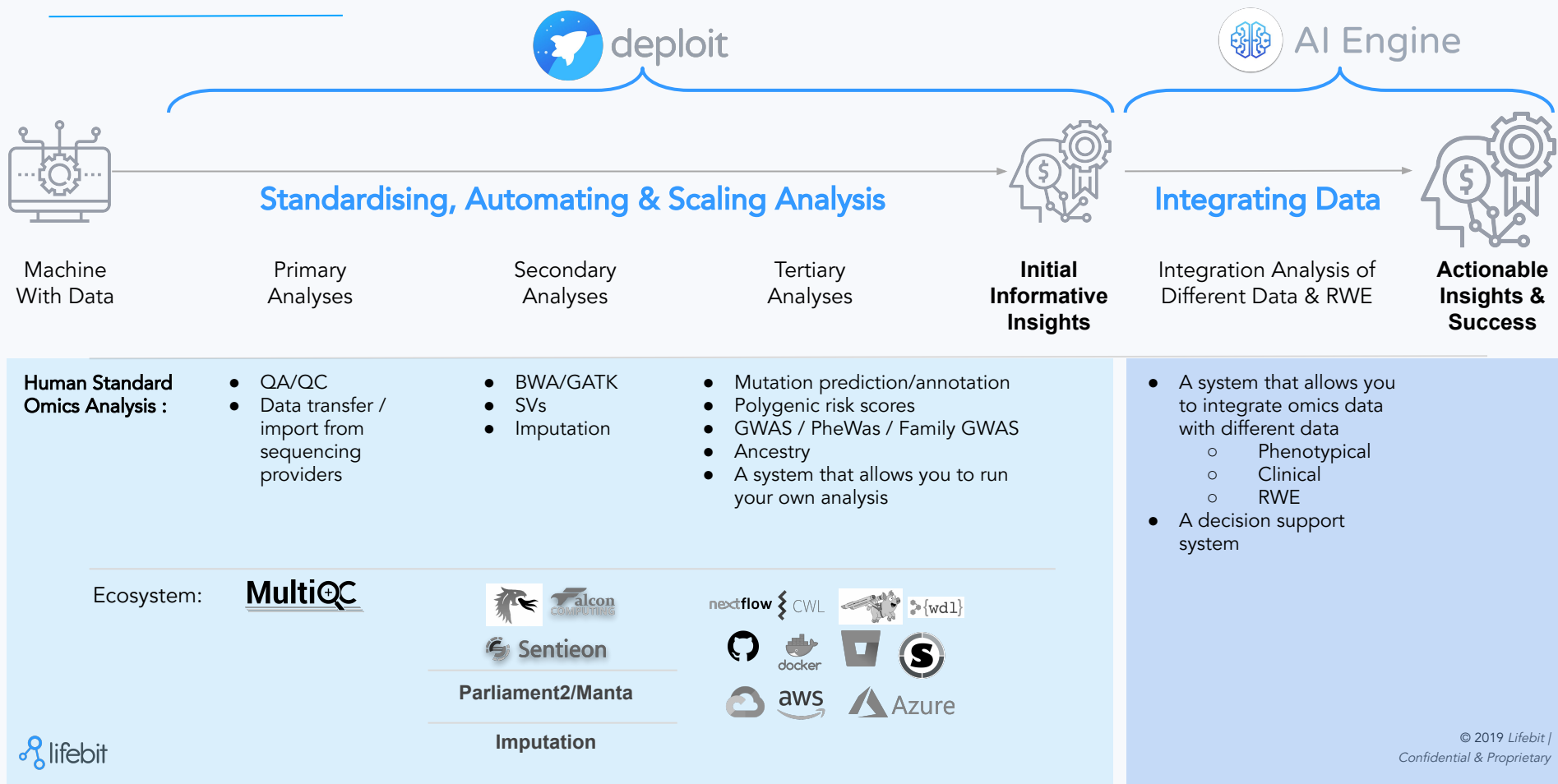




Welcome R-ladies!

Dr. Maria Chatzou
CEO & Founder, Lifebit

Unifying omic analysis and integrating data





To date over 25 companies & organisations trust Lifebit Deployit to:

1. Instantly standardise & scale their data analysis over your their cloud & HPC
2. Reduce costs by 80%
3. Never move data



Solving industry pain points

- **Wasting money** on cloud computing?
- Relying on **black box** omics platforms?
- Needing to **move data**?
- **Struggling to scale** analysis so that it is truly automated?
- **Missing out** on the latest & most innovative bioinformatic pipelines?
- **Lacking** a truly **hybrid** HPC & cloud setup?

 CLOUD COMPUTING

\$14.1 Billion in Cloud Spending to be Wasted in 2019



Katy Stalcup — January 3, 2019

 Follow @parkmycloud

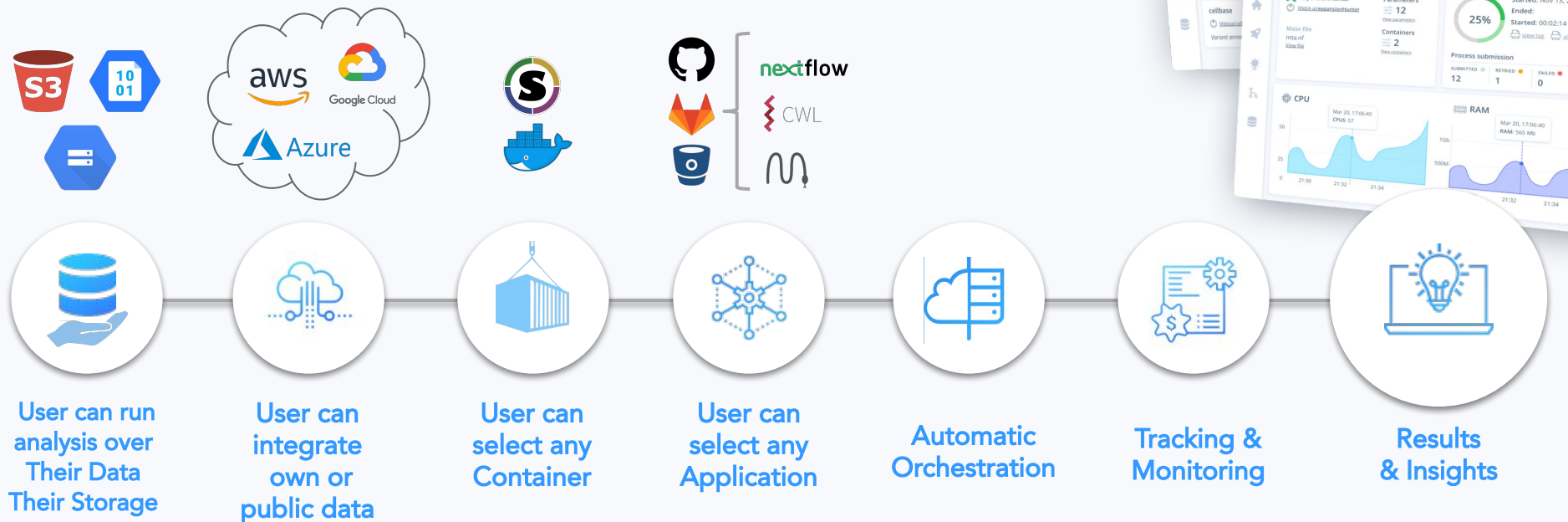
Big Pharma's bet on Big Data creates opportunities and risks

John Miller

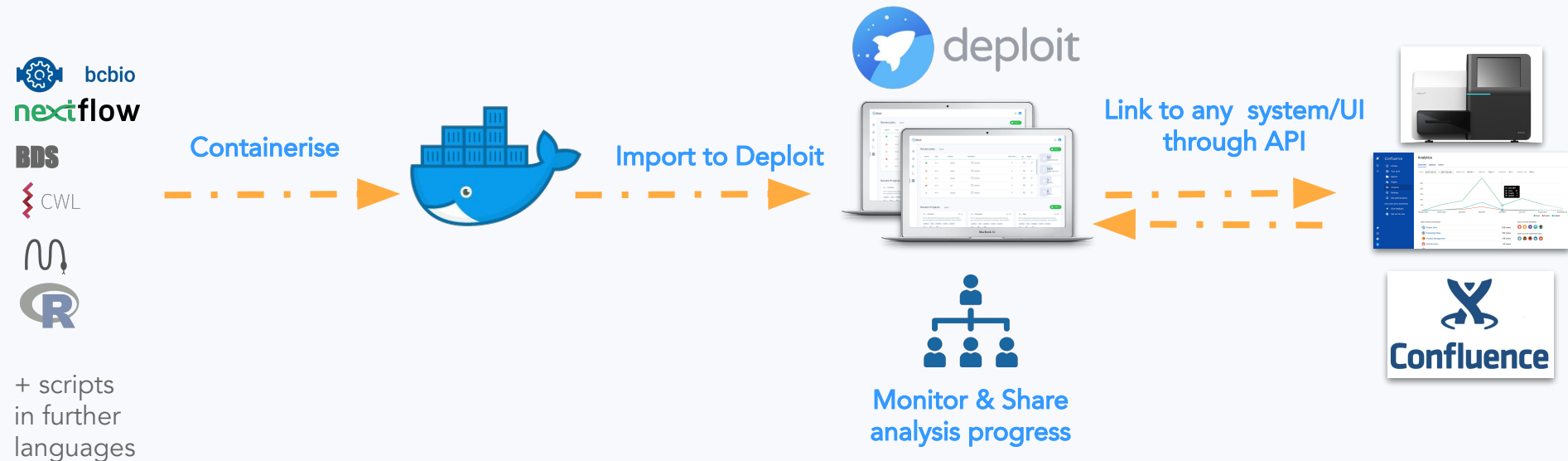
5 MIN READ



Orchestrating all analysis in one platform - in your environment



Instant adaption: Reuse & integrate every current pipeline



Run any existing pipelines on Deployt in 3 simple steps

Example: *Running QC pipeline for RNAseq data reusing existing bcbio pipelines*

Step 1: Dockerise the bcbio RNAseq pipeline

Step 2: Parse the Dockerised pipeline to Deployt by copy-pasting the link to the DockerHub repository

Step 3: Run the pipeline on Deployt using example open data

Port a new pipeline

* Docker Hub URL :

<https://hub.docker.com/r/lifebitai/rnademo>

Valid container!

* Name :

Bcbio RNAseq

Default command :

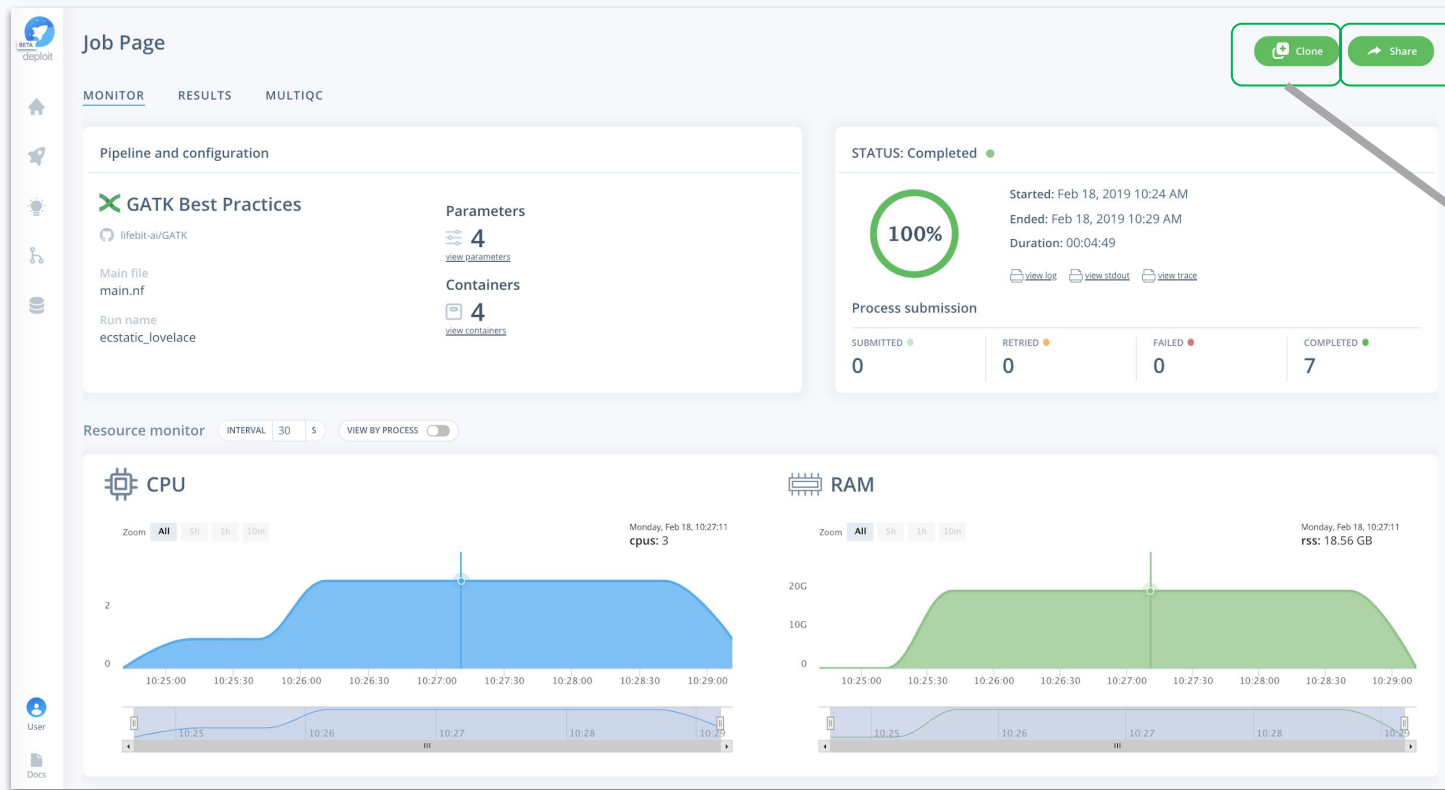
>_ /bin/bash /workflow.sh

Description :

bcbio RNAseq pipeline that uses Hisat2 aligner, Salmon for quantifications and DEXseq for Differential Expression Analysis

< Back Create pipeline

Run any existing pipelines on Deployt in 3 simple steps, **monitor**



share
the analysis

duplicate
the analysis

Run any existing pipelines on Deployt in 3 simple steps, monitor & get results

Job Page

MONITOR RESULTS MULTIQC

MultiQC

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

Report generated on 2018-12-21, 19:01 based on data in: `/Users/ewels/GitHub/MultiQC_website/public_html/examples/wgs`

General Statistics

Copy table Configure Columns Plot Showing 6/6 rows and 13/22 columns.

Sample Name	% GC	Ins. size	≥ 30X	Coverage	% Aligned	Change rate	Ts/Tv	M Variants	TiTV ratio (known)	TiTV i
P4107_1001	41%	358	74.7%	36.0X	97.3%	764	1.995	4.06	2.1	1.5
P4107_1002	41%	367	82.3%	40.0X	97.8%	762	1.994	4.07	2.1	1.5
P4107_1003	41%	365	82.4%	40.0X	97.6%	761	1.994	4.07	2.1	1.5
P4107_1004	41%	363	84.7%	46.0X	98.2%	765	1.996	4.05	2.1	1.5
P4107_1005	41%	368	85.3%	45.0X	98.0%	762	1.994	4.07	2.1	1.5
P4107_1006	41%	362	84.1%	43.0X	98.1%	761	1.993	4.07	2.1	1.5

QualiMap

Left sidebar menu:

- Snpeff
 - Variants by Genomic Region
 - Variant Effects by Impact
 - Variant Effects by Class
 - Variant Qualities
- GATK
 - Variant Counts
 - Compare Overlap
- Picard
- FastQ Screen
- 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 Quality Scores

Top right buttons: Clone, Share

Right sidebar: Toolbox, A, ?

Collaborate with your team in 2 simple steps

Add Users

* Add new user email and press enter :

+ Add

nick@gmail.com

x

Send Invites

1

Edit User · nick@gmail.com

Compute Cost Limit:
\$ 300

Access:
Member

Bucket	Read	Read/Write	No Access
astrazeneca-lifebit	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
benjamin-test-runs	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
blanpaindata	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
hellenic-pasteur-metagen	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
lb-az-data	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
lifebit-demo-apps	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
lifebit-partners	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
lifebit-user-data-007451c7-97d0-4c60-...	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
lifebit-user-data-05403b2c-e95c-46a9-...	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

Done

2

How we use R in Lifebit?

EDA & reporting

- 1 Set up misc
- 2 Explore metadata variables
- 3 Number of labels Group in study column
- 4 Number of labels Phenotype at sample collection column
- 5 Variables overview
- 6 Selected columns from the annotation sheet**

	Chip barcode	Chip Number	IonXpress barcode	Group in study	Phenotype at sample collection
1	DA0057156		001	Paired Norm-CIN3	Normal
2	DA0057156		003	Paired Norm-CIN3	CIN3
3	DA0059058	Chip 5	012	HPV pos control	Normal
4	DA0059045	Chip 1	005	Paired Norm-CIN3	Normal
5	DA0059045	Chip 1	006	Paired Norm-CIN3	CIN3

All

All

All

All

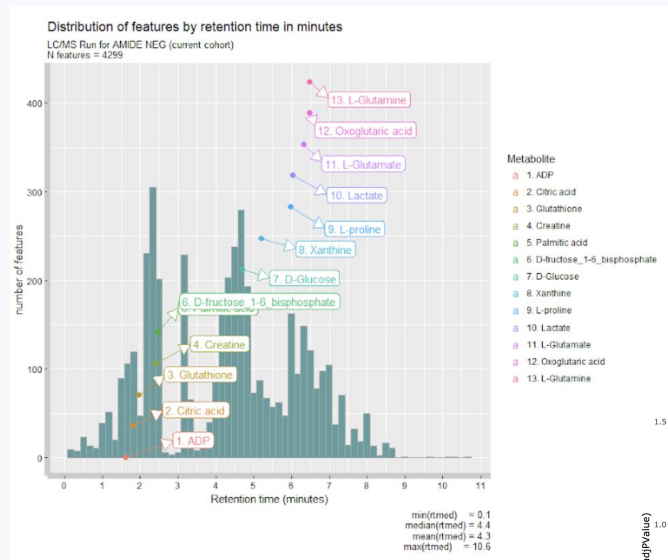
All

Showing 1 to 5 of 147 entries

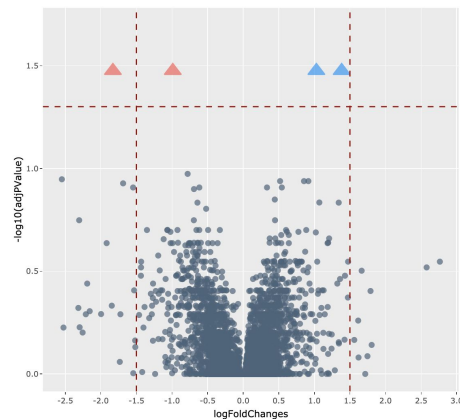
Previous 1 2 3 4 5 ... 30 Next

How we use R in Lifebit?

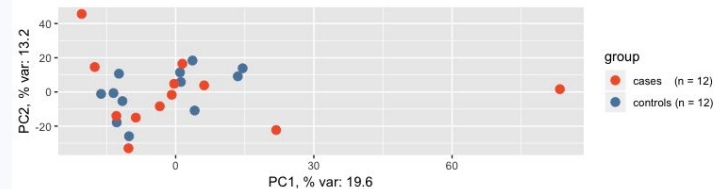
ggPlotting



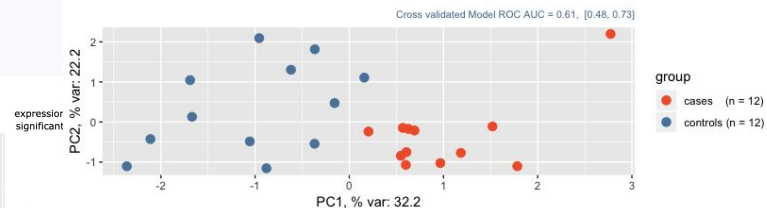
Volcano plot for genes in {CLL} ExpressionSet



PCA plot of Mesothelioma cohort (n = 24)
(Before feature selection)



PCA plot of Mesothelioma cohort (n = 24)
(After feature selection)



How we use R in Lifebit?

Blogging

TL;DR

Why call it 'fastQsee' though 🤔 ?

What will I learn?

Last things first: What results will I get from the FastQC tool on Deploy?

Overview

Step-By-Step Tutorial: Deploy in action (aka in gifs!)

Acknowledgements

fastQsee: A tutorial on how to use Docker containers on the **Deploy** platform to quickly generate a plot-ful FastQC report

Code ▾





Contact

Dr. Maria Chatzou
CEO & Founder

maria@lifebit.ai



Thorben Seeger
VP Commercial

thorben@lifebit.ai